

JOURNAL *of* EDUCATIONAL
RESEARCH

Volume No. 1, 1920

THE LIBRARY

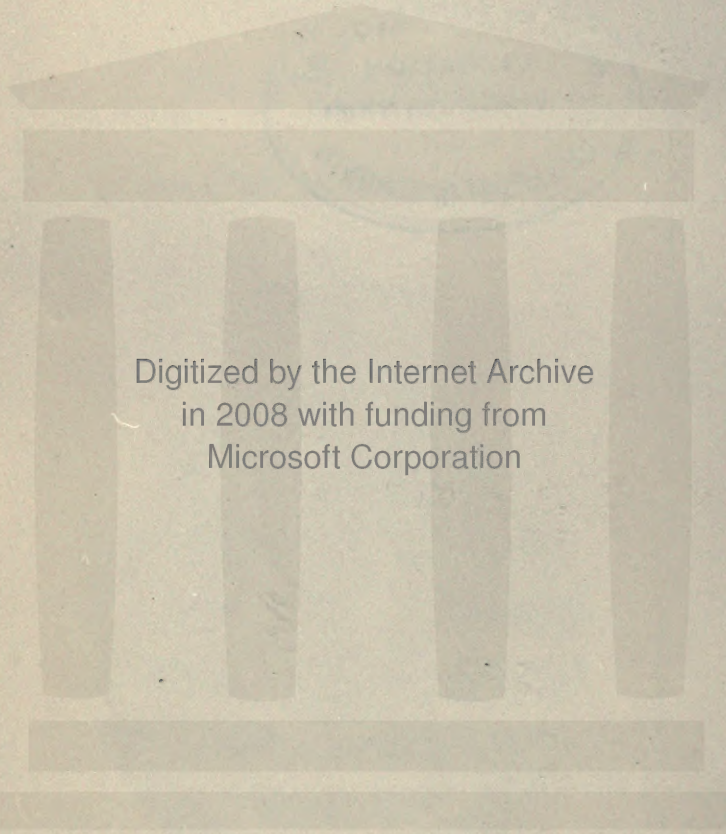
The Ontario Institute
for Studies in Education

Toronto, Canada



Presented by
Peter Sandford
to





Digitized by the Internet Archive
in 2008 with funding from
Microsoft Corporation

Journal of Educational Research

*Official Organ of the National Association of Directors of
Educational Research*

B. R. BUCKINGHAM, *Editor*

Associate Editors

E. J. ASHBAUGH

L. P. AYRES

W. W. CHARTERS

S. A. COURTIS

WALTER S. MONROE

GEORGE D. STRAYER

LEWIS M. TERMAN

VOLUME I

January-May, 1920

Published for the Bureau of Educational Research
University of Illinois
by
PUBLIC SCHOOL PUBLISHING COMPANY
Bloomington, Illinois

Copyright 1920 by
Public School Publishing Company

INDEX

Volume I, January-May, 1920

(E) indicates an editorial

(N) indicates a signed news item or communication

(R) indicates a book review

Achievements of the School Year (E), W. S. Monroe, 299

Age, Mental and Physical, in Relation to School Administration (E), B. R. Buckingham, 139

Algebra, a Series of Calibrated Objective Tests in, M. A. Dalman, 47

American Psychological Association, Account of Meeting (N), A. I. Gates, 415

Analysis of and Testing in Common Fractions, A. W. Kallom, 177

Application of Scientific Method in Evaluating the Subject Matter of Spellers, Clifford Woody, 119

Application to Tables of Distribution of a Shorter Method for Computing Coefficients of Correlation, L. P. Ayres, 295

Arithmetic, Analysis of and Testing in Common Fractions, A. W. Kallom, 177; Survey of the Social and Business Usage of, G. M. Wilson (R), 309

Ayres, Leonard P., Application to Tables of Distribution of a Shorter Method for Computing Coefficients of Correlation, 295; Coefficients of Regression, 398; Shorter Method for Computing the Coefficient of Correlation, 216

Ayres Spelling Scale Extended, 77

Barthelmess, Harriet M., Need to Teach Spelling of Derivatives (N), 322

Belting, P. E., High School Failures (F. P. Obrien) (R), 226

Bibliography of Standardized Tests for the High School, W. S. Monroe, 149, 229, 311

Bobbitt, F. M., Curriculum (R), W. W. Charters, 147

Brooks, Samuel S., Putting Standardized Tests and Scales to Practical Use in the Rural Schools, 392

Buckingham, B. R., Announcement of Journal of Educational Research, 1; Educational Diagnosis of Individual Pupils (C. A. Buckner) (R), 307; Intelligence of School Children (L. M. Terman) (R), 144; Mental and Physical Age in Relation to School Administration (E), 139; Proposed Index of Efficiency in Teaching United States History, 161; Scientific Curriculum Building (E), 404; Study in Educational Prognosis (E. K. Fretwell) (R), 307; Textbooks: Their Cost and Improvement (E), 222; Virginia Public Schools (Virginia Education Commission) (R), 407

Buckner, Chester A., Educational Diagnosis of Individual Pupils (R), B. R. Buckingham, 307

Building Survey, State-Wide, G. D. Strayer, 204

Bureaus of Educational Research, Operation of, W. W. Theisen, 382

Bureaus of Research, Organization of (E), S. A. Courtis, 69

Cameron, Edward H., Classroom Problems in the Education of Gifted Children (T. S. Henry) (R), 412

Charters, W. W., Constructing a Language and Grammar Scale, 249; Curriculum (F. M. Bobbitt)

- (R), 147; Revising the Course of Study (E), 403
- Classroom Problems in the Education of Gifted Children, T. S. Henry (R), 412
- Coefficients of Correlation, Application to Tables of Distribution of a Shorter Method for Computing, L. P. Ayres, 295; Shorter Method for Computing, L. P. Ayres, 216
- Coefficients of Regression, L. P. Ayres, 398
- College Entrance, Intelligence Examinations for, E. L. Thorndike, 329
- Commercial Education in the Public High Schools of the United States, L. S. Lyon (R), 64
- Connor, William L., A New Method of Rating Teachers, 338
- Constructing a Language and Grammar Scale, W. W. Charters, 249
- Correlation of Physical Health and Mental Efficiency, R. L. Sandwick, 199
- Course of Study, Revising (E), W. W. Charters, 403
- Courtis, Stuart A., Gary Public Schools: Measurement of Classroom Products (R), W. S. Monroe, 302; Health by Stunts (N. H. Pearl and H. E. Brown) (R), 147; Organization of Bureaus of Research (E), 69; Scientific Supervision (E), 67
- Courtis, S. A. and Packer, P. C., Educational Research, 5
- Curriculum, F. M. Bobbitt (R), 147
- Curriculum Building, Scientific (E), B. R. Buckingham, 404
- Dalman, Murray A., Hurdles, A Series of Calibrated Objective Tests in First-Year Algebra, 47
- Development of the Unit Plan of Penmanship Practice, H. C. Walker, 112
- Dickson, Virgil E., Gifted Child (E), 300
- Drawing, Talent in, H. T. Manuel (R), 409
- Educational, Diagnosis of Individual Pupils, C. A. Buckner (R), 307; Guidance of High-School Pupils, Use of Psychological Tests in, W. M. Proctor, 369; Legislation and Administration in the State of New York, 1777-1850, E. G. Hobson (R), 65; Measurements, Little Rock, Ark. (N), G. T. Huckaby, 243; Objectives of the Illinois High School Conference, H. A. Hollister, 129; Research, S. A. Courtis and P. C. Packer, 5; Tests as Important Schoolroom Equipment (E), W. S. Monroe, 68
- Elementary School, Self-Survey in, R. H. Lane, 172
- Encouragement of Individual Instruction by Means of Standardized Tests, M. H. Willing, 193
- Examination, a New Kind of, W. A. McCall, 33
- Failures, High School, F. P. Obrien (R), 226
- Fretwell, Elbert K., Study in Educational Prognosis (R), B. R. Buckingham, 307
- Gates, Arthur I., American Psychological Association—Account of Meeting, 415
- Gary Public Schools: Measurement of Classroom Products, S. A. Courtis (R), 302
- Gifted, Child (E), V. E. Dickson, 300; Children, Classroom Problems in the Education of, T. S. Henry (R), 412
- Grading School Children, Use of Intelligence Tests in, L. M. Terman, 20
- Grammar, Constructing a Language and Grammar Scale, W. W. Charters, 249

- Gray, William S., Educational Legislation and Administration in the State of New York from 1777 to 1850 (E. G. Hobson) (R), 65; Survey of Commercial Education in the Public High Schools of the United States (L. S. Lyon) (R), 64; Value of Informal Tests of Reading Accomplishment, 103
- Group Scale of Intelligence for Use in the First Three Grades: Its Validity and Reliability, L. W. Pressey, 285
- Hand Sewing, Measurement of Certain Elements of, Katharine Murdoch (R), 410
- Handwriting, Development of the Unit Plan of Practice, H. C. Walker, 112
- Health, and Mental Efficiency, Correlation of, R. L. Sandwick, 199; by Stunts, N. H. Pearl and H. E. Brown (R), 147
- Henmon, V. A. C., Improvement in School Subjects throughout the School Year, 81
- Henry, Theodore S., Classroom Problems in the Education of Gifted Children (R), E. H. Cameron, 412
- High School, Failures, F. P. O'Brien (R), 226; Pupils, Use of Psychological Tests in the Educational Guidance of, W. M. Proctor, 369
- History, Proposed Index of Efficiency in Teaching, B. R. Buckingham, 161
- Hobson, Elsie G., Educational Legislation and Administration in the State of New York from 1777 to 1850 (R), W. S. Gray, 65
- Hollister, H. A., Educational Objectives of the Illinois High School Conference, 129
- Huckaby, G. T., Educational Measurements, Little Rock, Ark. (N), 243
- Hurdles, a Series of Calibrated Objective Tests in First-Year Algebra, M. A. Dalman, 47
- Illinois High School Conference, Educational Objectives of, H. A. Hollister, 129
- Improvement, of Instruction through the Use of Educational Tests, W. S. Monroe, 96; in School Subjects throughout the School Year, V. A. C. Henmon, 81
- Intelligence, Examinations for College Entrance, E. L. Thorndike, 329; Tests in the Grading of School Children, Use of, L. M. Terman, 20; of School Children, L. M. Terman (R), 144
- International Educational Correspondence (N), T. E. Oliver, 153
- Journal of Educational Research, Announcement, B. R. Buckingham, 1
- Kallom, Arthur W., Analysis of and Testing in Common Fractions, 177; Reproduction as a Measure of Reading Ability, 359
- Kelly, F. J., General Report on Buildings and Grounds of Delaware (G. D. Strayer) (R), 225
- Kelly, F. J. and Loomis, A. K., Retardation in One-Room Rural Schools in Kansas, 271
- Lane, Robert H., Self-Survey in the Elementary School, 172
- Language and Grammar Scale, Constructing a, W. W. Charters, 249
- Loomis, A. K. and Kelly, F. J., Retardation in One-Room Rural Schools in Kansas, 271
- Lyon, Leverett S., Survey of Commercial Education in the Public High Schools of the United States (R), W. S. Gray, 64

- Manuel, Herschel T., Talent in Drawing (R), J. O. Powers, 409
- McCall, William A., New Kind of School Examination, 33
- Measurement of Certain Elements of Hand Sewing, Katharine Murdoch (R), 410
- Mental, and Physical Age in Relation to School Administration (E), B. R. Buckingham, 139; Efficiency and Health, Correlation of, R. L. Sandwick, 199
- Murdoch, Katharine, Measurement of Certain Elements of Hand Sewing (R), W. S. Monroe, 410
- Myers Mental Measure, 324
- Monroe, Walter S., Achievements of the School Year (E), 299; Bibliography of Standardized Tests for the High School, 149, 229, 311; Educational Tests as Important Schoolroom Equipment (E), 68; Gary Public Schools: Measurement of Classroom Products (S. A. Courtis) (R), 302; Improvement of Instruction through the Use of Educational Tests, 96; Measurement of Certain Elements of Hand Sewing (Katharine Murdoch) (R), 410; Survey of the Social and Business Usage of Arithmetic (G. M. Wilson) (R), 309; Teaching of Spelling (W. F. Tidyman) (R), 63; Value of Standardized Silent Reading Tests (E), 136
- National Association of Directors of Educational Research, 71, 158, 247, 325, 421; List of Members, 422
- National Intelligence Tests (N), R. M. Yerkes, 321
- Need to Teach Spelling of Derivatives, H. M. Barthelmess, 322
- New, Kind of School Examination, W. A. McCall, 33; Method of Rating Teachers, W. L. Connor, 338
- North Dakota Bureau of Educational Cooperation, 79
- O'Brien, Francis P., High School Failures (R), P. E. Belting, 226
- Oliver, Thomas E., International Educational Correspondence (N), 153
- Operation of Bureaus of Educational Research—with Particular Reference to Wisconsin, W. W. Theisen, 382
- Packer, P. C. and Courtis, S. A., Educational Research, 5
- Pearl, N. H. and Brown, H. E., Health by Stunts (R), S. A. Courtis, 147
- Powers, J. Orin, Talent in Drawing (H. T. Manuel) (R), 409
- Pressey, Luella W., Group Scale of Intelligence for Use in the First Three Grades: Its Validity and Reliability, 285
- Proctor, W. M., Psychological Tests as a Means of Measuring the Probable School Success of High-School Pupils, 258; Use of Psychological Tests in the Educational Guidance of High-School Pupils, 369
- Proposed Index of Efficiency in Teaching United States History, B. R. Buckingham, 161
- Psychological Tests, as a Means of Measuring the Probable School Success of High School Pupils, W. M. Proctor, 258; in the Educational Guidance of High-School Pupils, W. M. Proctor, 369
- Psychology of the Educational "Standpatter" (E), L. M. Terman, 135
- Publications Received, 65, 148, 228, 413
- Pupils, Psychological Tests as a Means of Measuring the Probable School Success of High School,

- W. M. Proctor, 258; Use of Psychological Tests in the Educational Guidance of High School, W. M. Proctor, 369
- Putting Standardized Tests and Scales to Practical Use in the Rural Schools, S. S. Brooks, 392
- Reading, Experiment in Silent, 414; Reproduction as a Measure of Ability in, A. W. Kallom, 359; Tests Used in Kane, Pa., 244; Value of Informal Tests of Reading Accomplishment, W. S. Gray, 103; Value of Standardized Silent Reading Tests (E), W. S. Monroe, 136
- Report on Buildings and Grounds of Delaware, G. D. Strayer (R), 225
- Reproduction as a Measure of Reading Ability, A. W. Kallom, 359
- Research and the Problems of Educational Readjustment (E), L. M. Terman, 137
- Retardation in One-Room Rural Schools in Kansas, F. J. Kelly and A. K. Loomis, 271
- Rural Schools, in Kansas, Retardation in One-Room, F. J. Kelly and A. K. Loomis, 271; Putting Standardized Tests and Scales to Practical Use in, S. S. Brooks, 392
- Salary Increases at Boston, 245
- Sandwick, R. L., Correlation of Physical Health and Mental Efficiency, 199
- School Administration, Mental and Physical Age in Relation to (E), B. R. Buckingham, 139
- Scientific, Curriculum Building (E), B. R. Buckingham, 404; Supervision (E), S. A. Courtis, 67
- Self-Survey in the Elementary School: an Experiment in Cooperation, R. H. Lane, 172
- Shorter Method for Computing the Coefficient of Correlation, L. P. Ayres, 216
- Southington Plymouth English Vocabulary Scale, 324
- Spellers, Scientific Method in Evaluating the Subject Matter of, Clifford Woody, 119
- Spelling, Need to Teach Spelling of Derivatives (N), H. M. Barthelmess, 322; Teaching, W. F. Tidyman (R), 63
- State-Wide School Building Survey, G. D. Strayer, 204
- Strayer, George D., State-Wide School Building Survey, 204; General Report on Buildings and Grounds of Delaware (R), F. J. Kelly, 225
- Study in Educational Prognosis, E. K. Fretwell (R), 307
- Supervision, Scientific (E), S. A. Courtis, 67
- Survey, of the Social and Business Usage of Arithmetic, G. M. Wilson (R), 309; Survey in the Elementary School: an Experiment in Cooperation, R. H. Lane, 172
- Talent in Drawing, H. T. Manuel (R), 409
- Teachers, New Method of Rating, W. L. Connor, 338
- Teaching of Spelling, W. F. Tidyman (R), 63
- Teaching, Proposed Index of Efficiency in Teaching United States History, B. R. Buckingham, 161
- Terman, Lewis M., Intelligence of School Children (R), B. R. Buckingham, 144; Research and the Problems of Educational Readjustment (E), 137; Psychology of the Educational "Standpatter" (E), 135; Use of Intelligence Tests in the Grading of School Children, 20
- Testing in and Analysis of Fractions, A. W. Kallom, 177
- Tests: Algebra, M. A. Dalman, 47; Educational Tests as Important

- Schoolroom Equipment (E), W. S. Monroe, 68; High School, Bibliography of, W. S. Monroe, 149, 229, 311; Improvement of Instruction through the Use of, W. S. Monroe, 96; Individual Instruction by Means of, M. H. Willing, 193; Intelligence Examinations for College Entrance, E. L. Thorndike, 329; Intelligence Tests in the Grading of School Children, L. M. Terman, 20; Intelligence Tests for Use in the First Three Grades, L. W. Pressey, 285; Language and Grammar, W. W. Charters, 249; Myers Mental Measure, 324; National Intelligence Tests (N), R. M. Yerkes, 321; Psychological, as a Means of Measuring the Probable School Success of High School Pupils, W. M. Proctor, 258; Standardized Tests in Rural Schools, S. S. Brooks, 392; Reading, A. W. Kallom, 359; Reading, Value of, W. S. Gray, 103; Reading, Kane, Pa., 244; Reading, Silent, Value of (E), W. S. Monroe, 136; Southington Plymouth English Vocabulary Scale, 324; State-Wide Use of in Colorado, 420; Virginia School Survey Tests, 324
- Textbooks, Their Cost and Improvement (E), B. R. Buckingham, 222
- Theisen, W. W., Operation of Bureaus of Educational Research—with Particular Reference to Wisconsin, 382
- Theses Listed, 79
- Thorndike, Edward L., Intelligence Examinations for College Entrance, 329
- Tidyman, Willard F., Teaching of Spelling (R), W. S. Monroe, 63
- United States Bureau of Education and Educational Research, 156
- Use of Psychological Tests in the Educational Guidance of High School Pupils, W. M. Proctor, 369
- Value of Standardized Silent Reading Tests (E), W. S. Monroe, 136
- Virginia Education Commission, Virginia Public Schools (R), B. R. Buckingham, 407
- Virginia Public Schools, Virginia Education Commission (R), 407
- Virginia School Survey Tests, 324
- Walker, H. C., Development of the Unit Plan of Penmanship Practice, 112
- Willing, Matthew H., Encouragement of Individual Instruction by Means of Standardized Tests, 193
- Wilson, Guy M., Survey of the Social and Business Usage of Arithmetic (R), W. S. Monroe, 309
- Woody, Clifford, Application of Scientific Method in Evaluating the Subject Matter of Spellers, 119
- Yerkes, Robert M., National Intelligence Tests (N), 321

Journal of Educational Research

Editorial Office, Urbana, Illinois
Editor-in-chief B. R. Buckingham, University of Illinois

Business Office, Bloomington, Illinois
Public School Publishing Company

VOLUME I

JANUARY, 1920

NUMBER 1

ANNOUNCEMENT

The educational public is entitled to know the reasons for establishing a new educational periodical, the place it is designed to occupy, and the purpose it is intended to serve. It is true that no assumption of a "mission" however happily expressed will weigh beside the revealing substance of the periodical itself. Like a plank in a political platform it will have merely the interest of suspense until practice squares with promise. To the hard-headed reader, it may even seem but an artifice to beguile him from his customary caution. Yet it is equally true that the possession of a clearly conceived plan may reasonably be required of those who thus presume to enter a field already so fully occupied.

We feel that we have such a plan. The JOURNAL OF EDUCATIONAL RESEARCH will be devoted to the uses which have been made and are capable of being made of the results and methods of investigation. It will emphasize applications rather than abstractions, and practice rather than theory. It will not avoid the theoretical altogether, and it will present some technical articles. But such articles will be most welcome when they have obvious practical implications. Research for the sake of research we shall leave for others. What uses the teacher and supervisor have made or may make of the findings of the experimentalist, what methods are transferable from the laboratory to the classroom and office, what workable means of meeting common needs have been developed under actual school conditions—these will be our chief concern. We shall seek contributions not only from

the leaders in research but also from their followers, the practical school people.

And the school people are likely to contribute handsomely. Current periodicals reveal that the number of them who have the true research spirit and the ability to report their findings effectively is increasing. New types of teacher-training have exalted a fact basis for instructional and administrative procedure; and the introduction of better school records and new means of measurement have made such a fact basis possible. What the facts are and what comes of them—the methods that work and the devices that succeed—may be told so that others may derive friendly lessons for their guidance.

Theory has outrun practice in our educational writings. Since scales and standardized tests began, ten years ago, to afford new instruments of precision, a new language has developed—a language whose vocabulary and idiom are strange. Statistics has come upon the scene, providing a powerful method of analysis, but at the same time assuming to a disquieting degree the possession on the part of the reader of knowledge and a type of thinking which he often lacks. Results are quite generally abstract, and “tentative conclusions” are often merely speculative. Some of us wonder whether all this is necessary—whether a writer may not be scientific without being impractical or profound without being obscure—whether the time has not come when some of these results may be restated in usable form—whether we do not now possess the materials for studying some problems long enough and hard enough to reach conclusions which are fundamental and satisfactory.

The attempt will, therefore, be to permit practice to catch up with theory. The Journal will endeavor to exemplify sound theory and it is hoped that its technic will be adequate. But we believe its theory and technic will be best exemplified when they are employed in the reporting of practical investigations and in the formation of practical results.

In its statistical presentations, the Journal will assay to set a standard for clear, simple treatment. Where unfamiliar terms or processes are used it will be our policy to define and explain them. As a part of this general plan a series of statistical articles will appear shortly which will attempt to deal in an

understandable way with the terms and processes in current usage in educational writings.

Although the JOURNAL OF EDUCATIONAL RESEARCH is published for the University of Illinois it is a national magazine. Its board of associate editors and its staff of contributing editors represent many sections of the country. Manuscripts are being accepted for publication and will continue to be accepted regardless of the affiliations of their authors. The National Association of Directors of Educational Research—an organization having bureaus in universities, in city school systems, and in one state department—has adopted the Journal as its official organ. These are evidences that the Journal will by no means be local in character.

Not only will it aim to appeal to educational people everywhere but it will also be hospitable to educational questions of every type. Indeed its scope in this respect will be as broad as education itself. For convenience, however, six major fields have been recognized, and corresponding to each is an associate editor.

Leonard P. Ayres will have charge of child accounting; W. W. Charters will give attention to curriculum analysis and derivation; S. A. Courtis will edit manuscripts relating to school supervision and the organization of bureaus of research; Walter S. Monroe will deal with educational tests and measurements; George D. Strayer will direct the policy of the magazine as to educational finance, buildings and building programs, records and reports; and Lewis M. Terman will have control of the department of intelligence testing. E. J. Ashbaugh, Secretary of the National Association of Directors of Educational Research, is also an associate editor. He will be responsible for the section of the magazine devoted to the affairs of the Association.

From one point of view no time could be less auspicious than the present for launching this venture. The scarcity and high price of paper combine with labor shortage and unrest in the printing and binding trades to produce a condition bordering upon chaos. Many magazines are suspending publication, and most others are limiting their output. Under these circum-

stances the editors and publishers cannot at present be as liberal as they would like to be.

From another point of view however the occasion is peculiarly favorable. There are unusual accumulations of good material seeking an outlet. On the one hand the facilities for publication have been scanty during the war, and on the other hand investigations have never been more numerous. Old prejudices have been abandoned. Precedent has meant less than ever before. A new sense of values has been created, and time-honored processes have been challenged. Some of them have stood the test and have been retained; others have been rejected; while still others have been modified and redirected. Those who have been engaged in this movement have something to say, and they have been unable to say it.

Whether under such circumstances a magazine having these avowed aims can meet the needs of the hour will depend not only upon the validity of the aims but also upon the steadfastness with which they are maintained. Believing that our reasons are good reasons, that the place we seek to occupy should be more fully occupied, and that our purposes are worthy, estimating the chances of success or failure, and convinced that success is certain if we keep faith with our intentions, we present to the educational public this first number of the JOURNAL OF EDUCATIONAL RESEARCH.

B. R. BUCKINGHAM.
Editor.

EDUCATIONAL RESEARCH
S. A. COURTIS AND P. C. PACKER

Public Schools, Detroit, Michigan

Many articles have appeared on the subject of educational research, but the movement for the use of scientific methods in education has grown so rapidly that already much of what has been written is out of date. It is fitting, therefore, that in the very first number of a journal devoted exclusively to the field, the attempt should be made to reformulate in terms of present day activities the functions of research in education, to define its scope, and to describe the organization of an ideal department of educational research in a large city school system.

The point of view of the authors of this article can be most readily understood if education is conceived as a war waged by society to gain control of its own evolution. Many years ago the human race, having won certain of the preliminary skirmishes against natural obstacles in earth and sky and sea, undertook to conserve its hard won gains by creating schools which should pass down to the next generation all that seemed to be of permanent value in the experience of the race. That is, at first education was a defensive warfare merely. Its chief goal was to conserve; to preserve unchanged.¹

With the development of human thought and philosophy has come a clearer vision. Today men understand as never before that the whole spirit and destiny of a people are molded and directed by the work done in the schools. Today in education society is beginning to conduct an aggressive warfare. Education now aims not only to conserve the best of the past, but to improve upon the present and to provide for a progressive future. Today the discovery and exact description of the instincts, capacities, and tendencies of children, the invention of more efficient methods of instruction, the use of more certain methods of direction, and the achievement of better forms of organization of human energy for carrying forward the collective struggle for world betterment are seen by many to be the supreme and persistent battle in man's ever widening and progressive conquest of nature.

¹ Well illustrated in China under the old régime.

But nature is perverse. In each age she marshals her forces anew. Old instincts ever tend to find lawless expression under each set of new conditions. Just as the Hollander must repair his dikes constantly to keep the sea back from the lands already reclaimed, just as further gains make greater and greater demands upon engineering skill and knowledge, so each age has its own educational adjustments to make. The educational warfare is a real war. By frontal attack, by surprise raids from air and sea, by sapping and mining, the struggle goes constantly on. If one age breaks down a feudal system which blocks the way to progress, the next develops factory production and industrial slavery. The nations have gone to war to make the world safe for democracy, but if the schools do not raise a breed of men who will put democracy into industry and at the same time make a wise use of the increase in the wealth and leisure time that will result, the sacrifices of civilization are likely to have been in vain. The true battle front of progress in increasing measure is to be sought in our schools.

If now we ask ourselves how a successful general goes about *it to win his battle*, we find that one very important element to be considered is the efficiency of his research department. His advance listening posts, his scouting parties, his aeroplanes, maintain a continuous survey of the enemy and of all his activities. The general must know his opponent's strength and the disposition of his forces, he must watch his movements and anticipate his attacks. His success depends upon his ability to detect and interpret every change in the situation, whether that change spells trouble or opportunity.

Let us call this watchful, continuous collection of facts the "survey" function of the intelligence department. It is the first of the two great functions of all research.

It is, however, by no means the only function. For instance, suppose the survey work of the intelligence department of an army is well done. That is, suppose the commanding officer has at his command reliable and reasonably complete information in regard to his opponent, and decides upon offensive action. He makes his plans, and his men and guns move forward. Immediately a new type of research work begins. It is now the supreme duty of his intelligence department to keep the general in

constant touch with what happens on every front. His plans fail here, succeed there. At one part of the line his armies move over into new ground and new conditions. At another, the enemy develops unexpected strength. If there is to be eventual success, the general must be able to modify his plans on the basis of the progress of the battle. He must be able to judge at every instant how well he is succeeding. Coordination of effort and consolidation of gains are directly dependent upon this "judgment" activity, and this in turn upon the efficiency of his research department.

Let us call this second type of service the appraisal function of research. In its essential nature all research, as the name implies, is simply the organized, careful, and systematic search for facts, but it is well to emphasize the different purposes for which research may be carried on. "Survey" and "appraisal" merely stress the "before" and "after" aspects of the relation of research to executive action. No matter how great the executive genius which plans, or administers, achievement will ever, except by chance, be conditioned, on the one hand, by the completeness and accuracy of the information upon which plans are based and, on the other, by the extent to which they are modified to fit actual conditions as these are revealed through executive action.

If now we turn from the battle field to the classroom, the exactness of the parallel must be apparent. The superintendent as executive head of the school system is constantly called upon to plan and to execute. He has need to know the facts; he has even greater need to know the effects of his decisions. In educational warfare, we are as yet fighting blindly and in the dark. The intelligence service is not well organized. Too often apparent success at one part of the line appears later as a detriment to the general progress. For enduring success there must be active and continuous survey work and even more active and more persistent appraisal of the effects of executive action.

That is, an ideal superintendent in an ideal school system would have as two of his fundamental rules of action, the following:

1. Except under the pressure of the need for immediate action I will never plan an executive

policy or action without a preliminary survey to secure the facts which will make my planning intelligent, and

2. Will never perform an executive act without immediately taking steps to collect the facts which enable me to appraise the effect of the act.

In such an ideal school system under such an ideal superintendent, there would be need for a large department of educational research and the department would be kept very busy.

It is not enough, however, to have an intelligent appreciation of the two functions of research. Every school man needs also a deep-seated conviction as to their importance. Perhaps this can be gained best from an illustration from biology.

In the study of any living organism from a one-celled animal to man himself, two types of nervous mechanisms are found; the sensory and the motor. The first transmits inward the stimulations derived from the environment; the other transmits outward to the muscles those stimulations which result in motor activity.

Take the case of a cat, for instance. Suppose the controlling intelligence of a cat receives from its department of the interior a report that supplies are running low; in other words, that it is hungry. Suppose further that the cat then directs its research department to survey the landscape to secure the data upon which to base a plan for obtaining the desired food, and that soon the nose reports that a mouse is smelled, the ears that a mouse is heard, the eyes that a mouse is seen.

The cat is now in possession of the necessary facts. She formulates a policy; she will catch the mouse. The executive orders are issued, the muscles contract, the cat jumps for her prey.

The research department now begins to perform the appraisal functions. The nose and ears and eyes measure the distance and report that the legs jumped short; that further action is necessary if the goal is to be achieved.

In the meantime, let us suppose that the mouse has been doing some surveying on his own account and decides that he has important business elsewhere. His executive actions are at once

reported by the cat's research agents. Her original plan is then modified, the length and rapidity of her jumps are changed until the mouse is caught.

Returning to the educational situation once more, we find that in the original one-room private school in early colonial days, the master of the school performed every educational function. He gathered his own facts, he formulated his own policies, he acted upon them, and he appraised the effect of his actions, even as many one-room school teachers do to this day. But as public education grows to be a great municipal, or state, or national system, this is no longer possible. There has to be division of labor and specialization. Today the technic of research demands special training and undivided attention, both of which are seldom possible for the average superintendent. Research functions must be delegated to a special department.

An important question to be considered is the magnitude and importance of the work of such a department. The answer is to be found in our biological parallel. For in nature there have been two types of development: plants and animals. In most plants, the sensory or research functions have undergone almost no development from those of the original cell, while the motor or executive functions make up the main life activities of the plant. Thus a tree is a large and complex organism which carries on energetically for many years activities of surprising magnitude, yet its sensory mechanism is but little more developed than that of the elemental cell from which it has evolved.

In the animal kingdoms, however, the development of the sensory mechanism has paralleled the complexity of the structural and motor elements. In animals, life itself depends upon the proper coördination of the functioning of the two types of mechanisms. The plant organization is admirable as a permanent adjustment to a fixed environment, but the animal organization permits greater freedom, greater range of action, and thus makes possible the attainment of higher levels of progress.

If now we examine a large school system, such as that of New York City, which spends in one year \$45,000,000 on its schools, and whose superintendent is called upon to direct the activities of an army of 35,000 teachers and 900,000 children, and if we ask ourselves which type of organization it most

resembles, the judgment must be that it is the vegetative type. It is the executive, the administrative organization which has developed, not the research or sensory mechanism. The survey and appraisal activities upon which executive actions are based are carried on today as in the beginning, on the basis of opinion merely. It is true a beginning has been made. Since 1912 there has been a department of reference and research, but the total amount spent in New York City during the school year 1918-1919 for research was less than a tenth of one per cent of the total budget. In no sense does either survey of appraisal play a vital part in determining executive action.

What is true of New York City is true also of our educational life in general. We have developed education as a vegetative and permanent adjustment to a fixed environment, when all our experience and all our reasoning proclaim in stentorian tones that the environment is not fixed, that the present adjustment of education to society is not permanent, that the needs of this day and of every day to come is for progress, progress, and yet more progress. In other words, the superintendent of any of our large city systems has about as much chance of making his schools serve adequately the purposes in society which we now see they ought to serve, as a cat would have of catching a mouse if its research mechanism were dissected out and replaced by the elementary sensory machinery of some of the lower animals from which the cat has evolved.

The conclusion is clear and cannot be avoided. Educational research is destined to develop until it parallels in magnitude and importance the administrative phase of education. Whenever the activities of any department attain such size and complexity that there must be division of labor, the creation of every administrative agent must be paralleled by the creation of a corresponding research agent, except where the two functions can be combined in a single person. Thus the superintendent will necessarily combine within himself both the research and the executive functions. He must be able to initiate, direct and understand research work and methods in every field, as well as to formulate policies and administer the same in every field. That is, the ideal superintendent will be as much the director of educational research in his school system as he is now the chief execu-

tive officer. And there is probably no more fitting training for him who aspires to be superintendent than several years of service in the various fields of research.

Few school men yet realize that educational research means something more than the mere giving of a few trivial standard tests in the grades; that education will be a science only when scientific attitudes and scientific methods pervade every phase of educational work. Perhaps it would be more nearly correct to say that few have ever thought of education except in the narrow terms of their own special work. It will come as a surprise to many that in a fully developed school system, educational research would be carried on by some fourteen departments or divisions, each with a responsible head and each with research or clerical assistants. Several of the departments would need a large staff of specialists.

If we attempt to list these departments in logical order we must begin with (1) the department of sociological research. The superintendent needs an officer and a properly equipped department whose sole duty it is to survey society and report its structure, its condition, its tendencies, its needs. At present social pressure must reach high tension before it operates with certainty through the school board and the public press to influence the activities of the superintendent. In the future the schools must lead, not follow. Educational policies must be based upon adequate knowledge of the social situation, past, present, and future. Nor is it enough that the superintendent make such adjustments in organization, studies and methods of teaching as seem to him to produce the right effect upon society. There must be an accurate and scientific appraisal of the effect of the school's output upon society.² It is to this sociological appraisal that we must look for the final word in regard to every change made in the educational process. Survey and appraisal in the field of educational sociology are thus the alpha and the omega of education itself.³

² Meriam, J. L. "Measuring school work in terms of life out of school," *School and Society*, 5: 339-42, March 24, 1917.

³ Hanus, P. H., and Gaylord, H. D. "Courtis arithmetic tests applied to employees in business houses," *Educational Administration and Supervision*, 3: 505-20, November, 1917.

The second department would study (2) the methods of listing and keeping track of the children to be educated (census) and of getting them to school and of keeping them there as long as their presence in school is profitable. Few cities have any adequate knowledge of where or under what social conditions their children live.⁴ Still fewer make such studies of social changes that they can predict the population shifts in even the immediate future. Yet it is upon such data alone that wise purchases of school grounds can be made, or an adequate building program framed. Similarly, school officers have kept attendance records for years, but the methods employed have been so imperfect, and so little attention has been paid to the results that even today few adequate attendance data exist. Most school men, as well as the general public, would be astounded if the real facts in regard to absences, transfers, and removals could be presented in graphic form. Still less has been done in determining causes of variable attendance and in devising corrective measures. Attendance research is a wide field that reaches from the home and society on the one hand to the classroom and the teacher on the other. Upon the solution of certain of the problems in this field, depends directly the efficiency of work in several other departments.

The third field of research covers (3) the problems of organization and administration. Because of our present ignorance, for many years to come any scheme of organization must, at best, be but a compromise between desirable but conflicting ends. In the absence of scientific, experimental data, convention is usually the determining factor. Changes come slowly. Yet change, progress in educational ideals, demand at once corresponding change in forms and methods of organization. These in turn react upon every form of educational activity.⁵ At present the research bureaus of many cities are struggling with the problem of appraising such recent adoptions as the platoon organization, sectioning on the basis of intelligence tests, and the like. Properly functioning departments of administrative research would

⁴Ayres, L. P. *Child accounting in the public schools*. (Cleveland Education Survey.) Cleveland: Survey Committee of Cleveland Foundation, 1915.

⁵Strayer, G. D. and Bachman, F. P. *The Gary Public Schools: Organization and Administration*. New York: General Education Board, 1918.

have supplied experimental data to justify the form of organization adopted and to guide the efforts at appraisal.

The adoption of a definite educational policy and a definite form of organization would seem to demand (4) buildings and grounds to correspond. Yet research in this field is little more than an idea, a dream.⁶ At Gary, an educational ideal determined the form and arrangement of the building and grounds, but exceptional as this is in American education, few data can be shown which are a real aid in determining the best size, shape and arrangement of rooms, halls, etc., to achieve a given educational objective. However, a beginning has been made here also and standards for measuring existing buildings (an appraisal activity) should lead directly to architectural experimentation. A department of architectural research will be the eventual outcome, and a member of the research staff to devote his whole time to this field is immediately essential to an adequate department of research in a large city school system.

It is not enough to build schoolhouses educationally adapted to a given program. They must be kept in proper condition. Lighting, cleaning, ventilating, heating, repairing, etc., condition use as much as structure, perhaps more.⁷ Yet the attention given such details in most school systems is of the crudest sort. As educational research expands, as the search for the conditioning factors of educational efficiency is pushed farther and farther, it is inevitable that the day will come when a special department will be created to deal with experimental work of this particular kind. Let it be called (5) a research department for maintenance activities and janitorial service.⁸

⁶ No high school building has been designed about the educational activities that are to go within it except in the roughest sort of way. For instance, a high school building in one of the large cities of the country, constructed within the last three years, has lavatory space adequate for 2,400-2,600 pupils, but the shortage of classrooms is so great the building is overcrowded with 2,000 students.

⁷ Brkowitz, J. H. *Standardization of medical inspection facilities*. (United States Bureau of Education Bulletin No. 2, 1919.)

⁸ Thorndike, E. L. and Ruger, T. J. "The effects of outside air and recirculated air upon intellectual achievement and improvement of school pupils: a second experiment," *School and Society*, 4: 260-64, August 12, 1916.

Closely connected with buildings and maintenance are (6) problems centering around equipment and supplies. The determination of what equipment should be purchased, what supplies should be used, belongs primarily to the department of instructional research, but once these questions are decided there arise problems of purchasing the needed supplies and equipment, of distributing them through the system, and of keeping them in good repair. Here again the plans of the best of superintendents and the efforts of the best of teachers and supervisors are often seriously handicapped by the failure to give to problems of purchase, distribution, and maintenance, the educational attention they deserve. Efficiency here does not mean the purchase of the cheapest material from the lowest bidder, nor the reduction of the expense of distribution or maintenance to the lowest possible level. Only the research man whose business it is to trace evident results to their fundamental causes can appreciate the number and importance of the research problems in this field. They demand the existence of a divisional organization to deal with them.⁹

A group of research problems whose importance is being recognized more and more widely each year are (7) those which have to do with the selection, classification, assignment, and promotion of both the personnel of the educational corps, and of the children.¹⁰ As education becomes more and more scientific every educational agent, from cleaning women to teachers, from stenographers to superintendent, is chosen more and more on the basis of fitness for his particular job. Fitness is measurable and careful selection by appropriate tests saves the waste and disaster caused by the unfit. "Promotion on the basis of merit" presupposes that merit can be determined in some more objective and impersonal way than by mere opinion. As yet, however, scarcely a beginning has been made. Anyone can enumerate more problems which press for solution than the wisest man could solve in many years of work.

⁹ Boston, Mass: Finance commission. Report on the Boston School System, 1911.

¹⁰ Saam, Theodor. "Intelligence testing as an aid to supervision," *Elementary School Journal*, 20: 26-32, September, 1919.

Even more clearly recognized is the need for careful determination of the original nature of children. It is evident that for the protection of the school, as well as of the child himself, each new scholar must be given a searching diagnostic physical, mental, educational, moral, and social examination, immediately upon his entrance into the school system and that these examinations must be repeated at sufficiently short intervals to detect causes of trouble before they have a chance to produce serious effects. The proper classification, assignment, and promotion of children are functions so important in education that they call for a large and complex organization to perform them in any adequate manner. The inefficiency of our present machinery is daily becoming more apparent.

It is only a question of time before the same department will be extended to include the examination, selection, assignment, and promotion of teachers, janitors, and every other type of educational workers. Not every person who completes successfully a normal training course is fit to be a teacher. Indeed little is known about what elements of training and character determine teaching ability, while the selection of janitors, engineers and other persons whose work has to do with the operation of schools has scarcely been recognized as yet as an educational problem. In many schools the ideals and refinement of classroom work are daily being offset by contacts in basement and playground with vulgar and immoral janitors and engineers. Each of these has a particular work to do, and the work can be done efficiently only by some one who is both trained and adapted to the work he has to do. Research problems of the greatest importance to society center in this field.

One of the largest and most complex divisions is (8) the department of instructional research, and many are the types of problems which would occupy the research men specializing in this field. Wherever in the school system there is classroom work, there will be found problems of standards, courses of study, comparative efficiencies of different types of textbooks, equipments, and methods of work. Measurement of the effects of supervision is almost a field in itself, as are also the problems arising in normal training, teacher training, technical training,

etc., the education of typical children, the brilliant, the blind, the deaf, the crippled, the mentally deficient, and all the other types of children who depart widely from the average child. Every one of these fields has its peculiar problems which can be solved only by one who is thoroughly familiar with the special technic of the field, yet the instructional element is a common factor tying all together into a single department.

Another important division of research is that which has to do with (9) problems of health education. It is known definitely that physical health conditions mental efficiency, but the American public is just awakening to the fact that the health habits of children, the hours and conditions of sleep, the amount and quality of food, the amount of healthful recreational exercise, are vital elements in educational efficiency and must sooner or later be brought under social control. In the meantime the entire range of problems covering the relation between physical and mental ability, and physical and mental growth, offers almost a virgin field in research!

Still another distinct field of research is to be found in (10) cost accounting and the problems of educational finance. Research problems of the proper relation between taxation and public wealth of relative costs of administration, supervision, and teaching, of optimum versus minimum values, and a host of similar questions press for solution.¹¹

Mention must also be made of the research problems found in (11) vocational guidance and continuation schools; of those in (12) adult education and Americanization work just now in the public eye; of those connected with the work of (13) preserving suitable records of the achievements of the school system, of preparing reports and statistics for publication, and of conducting publicity campaigns for the enlightenment and education of the community. Finally there must be (14) pure, scientific research, the development of new technical methods of measurement, new tests, new devices, and new organization of existing science. There must be the search for the natural laws under-

¹¹ Bachman, F. P. and Bowman, Ralph. *The Gary Public Schools: Costs for School Year 1915-1916*. New York: General Education Board, 1918.

lying the teaching-learning process, and there must be also the appraisal of the effect of all the research work outlined above.

One of the problems growing out of the development of research departments is that of the relation of such departments to the school board, to the superintendent, and to other school departments. It is evident that primarily research departments are merely special organizations to which the superintendent has delegated specific powers and functions and that they are acting for him and under his direction and control. However, in as much as the superintendent is obligated to make a report to his board from time to time in regard to conditions in the school system, it should be evident that all possibility of manipulation of the results of various research activities should be eliminated as far as possible. That is, the research department, on request, should report directly to the board. This, however, means divided responsibility, while the whole tendency of modern thinking in administrative methods is towards fixing responsibility. The solution of the difficulty is not apparent, but one thing is certain; as research methods gain in effectiveness, control would tend to pass more and more into the hands of the research men. That is, in the future, an essential qualification of an executive officer in any field will be that he have the point of view which comes from a thorough grounding in scientific methods and that he be able to organize, direct, and utilize, research activities in his field. If his work becomes so heavy that part of it must be delegated, it is the research functions that will be retained. Cheap clerical assistants can attend to the details of administration if they are given a minimum of oversight and direction, but to recognize a problem before it has been recognized by others, to frame a policy for dealing adequately with the situation, and to appraise the effect of one's policy, calls for constructive, creative, effort of a high order.

Scientific measurement of educational products is about ten years old, but educational research in its broadest sense is as old as education itself. Nevertheless, all forms have shared in the increased vigor and effectiveness which have come from the rapid development and statistical technique of instructional research. Organized research departments in the modern sense

may be said to date from the Report of the Committee on School Inquiry, New York City, in 1912. Today there are forty or more bureaus of research in university, state or city systems, and in scores of small schools much research work is carried on by superintendents or other school officials. It may be instructive to ask how far educational research in any system approximates the ideal described above.

The city of Detroit established a Department of Educational Research in September, 1914, but the person placed in charge was interested primarily in a single phase of research work—namely, instructional research—and the activities of the department were restricted as much as possible to this field. Instructional research, however, leads to a study of administrative details, and the needs of a big city system produce a steady pressure for research in every field of work. The development of the research staff in Detroit has thus been an erratic and chance evolution, rather than the result of careful planning. Nevertheless, in September, 1919, the research forces of the city as represented by men specially trained in research methods were:

1. An assistant superintendent in charge of administrative research (buildings and grounds, budget making, organization, etc.)

2. A director of educational research in charge of supervisory and instructional research and miscellaneous research problems.

3. An assistant director of educational research in charge of research problems in secondary education.

4. A member of the normal school faculty in charge of research problems related to normal training and the training of teachers in service.

5. A director¹² of the department of special education in charge of mental testing, the psychological clinic, and related problems having to do with classification, assignment and promotion.

Each of these men has clerical and other assistants and each serves to direct the experiments and research problems of super-

¹²Position unfilled at the opening of school, but the work carried on by other members of the department.

visors, principals, and teachers throughout the city. In addition there are other departments which do much research work, as the department of records, publicity and statistics, while the Board of Education itself is organizing and directing a general survey of school work which makes use of and supplements the regular research activities.

The total volume of research throughout the city is thus very large. In the formal measurement of educational products alone the program calls for the use of over a million test papers during the year, while the informal and practice tests in use in the city will amount to half a million sheets more. All this means that in five years the movement for scientific methods has so permeated the system and touched the activities of educational workers at so very many points that the ideal organization and plans sketched above does not lie in as remote a future as the conventional school man of the old type is likely to imagine. Therefore, the chief purpose of the JOURNAL OF EDUCATIONAL RESEARCH is to bring to school men generally such a knowledge of the practical workings of educational research in the average school as will hasten the day when the ideal shall have become common practice.

THE USE OF INTELLIGENCE TESTS IN THE GRADING OF SCHOOL CHILDREN

LEWIS M. TERMAN

Leland Stanford Junior University

The early experiments with intelligence tests were directed chiefly toward the better understanding of the feeble-minded, and even their later development has centered largely around their use with defectives, delinquents, and other types of mental deviates. It is becoming clear, however, that their greatest usefulness will be found in their universal application to school children. Binet's "methode de luxe" has been so simplified that any intelligent teacher can use it in a way to enlarge greatly her knowledge of a given child. The group examination methods now in process of development may be expected to add still further to the popularity of intelligence tests as an aid in the more accurate grading and in the wiser educational guidance of school children. "A mental test for every child" is no longer an unreasonable slogan.

THE EXTENT OF INDIVIDUAL DIFFERENCES IN INTELLIGENCE

Intelligence tests have thrown into bold relief the enormous individual differences in intelligence which obtain for any unselected group of children of a given age. These differences are no less striking than those long known to exist in the case of height, weight, and other physical traits. The following figures illustrate typical individual differences in height, in ability to win promotion in school, and in intelligence, and show how similar the extent and frequency of the differences are in the three respects.

TABLE I. HEIGHT IN C. M., 236 BOYS, AGED $9\frac{1}{2}$ TO $10\frac{1}{2}$ YEARS.
(BALDWIN)

| Height in C.M. | 119-123 | 124-128 | 129-133 | 134-138 | 139-143 | 144-148 | 149-153 |
|---------------------|---------|---------|---------|---------|---------|---------|---------|
| Percent..... | 2.1 | 11 | 22 | 29.2 | 23.7 | 10.2 | 1.7 |

TABLE II. GRADE LOCATION OF 1896 UNSELECTED TEN-YEAR-OLD
BOYS, SALT LAKE CITY

| Grade..... | I | II | III | IV | V | VI | VII |
|--------------|-----|-----|-----|------|------|-----|-----|
| Percent..... | 0.7 | 4.8 | 24 | 49.4 | 18.6 | 2.3 | 0.2 |

TABLE III. I. Q.'S OF 83 UNSELECTED TWELVE-YEAR-OLD CHILDREN ^a

| I. Q. | 66-75 | 76-85 | 86-95 | 96-105 | 106-115 | 116-125 | 126-up. |
|--------------|-------|-------|-------|--------|---------|---------|---------|
| Percent..... | 5 | 15 | 20.5 | 28 | 19.5 | 11 | 0.8 |

^aI. Q. is the usual abbreviation for intelligence quotient, which is the quotient of mental age divided by actual age.

TABLE IV. MENTAL AGES OF 263 ELEVEN-YEAR-OLD CHILDREN

| Mental Age..... | 7 or less | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 or more |
|-----------------|-----------|------|------|-------|-------|-------|-------|------|------|------------|
| Percent..... | 2.66 | 5.32 | 9.50 | 15.96 | 22.81 | 14.07 | 12.93 | 8.36 | 5.32 | 3.04 |

Since the grade of school work which a child is able to do depends chiefly upon the level of mental development he has attained, such individual differences as those just noted for intelligence show the weakness of any method of school administration which tends to advance the pupil on the basis of chronological age. The chronologically old and the chronologically young may and often do belong together; the mentally old and the mentally young do not. Investigations show that notwithstanding the shifting which takes place at the end of each school year, the resulting classification of children has been so far from successful that, generally speaking, the lowest 20 or 25 percent of pupils in any grade belong mentally in a lower grade and the highest 20 or 25 percent in a higher grade. Only the middle 50 or 60 percent are classified approximately where they should be. Usually more than 15 percent of pupils are at least two grades removed from the one in which they belong by mental age.

The one criterion of fitness for promotion should be ability to meet the requirements of the next higher grade. Actually this criterion is lost sight of. The so-called "retarded" children are in reality usually from one to three grades above where they belong by mental development; the real retardates are the under-aged children who are generally found from one to three grades below the location which their mental development would warrant. The dull child is kept at tasks which are hopelessly beyond

his ability; the child of superior intelligence lacks the mental and moral stimulus which comes from intense application to tasks commensurate with his ability. In other words, the problem of retardation is exactly the reverse of what it is popularly supposed to be.

The failure of the school to grade children according to mental ability is shown by the overlapping of mental ages in the different grades. The following figures for certain first grade, fifth grade, and first year high school classes are typical.

TABLE V. MENTAL AGE (3-3 TO 3-11, ETC.)

| Grade | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 |
|------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| First..... | 4 | 22 | 31 | 51 | 32 | 6 | 2 | 1 | | | | | | | | | |
| Fifth..... | | | | | 1 | 1 | 10 | 26 | 17 | 13 | 6 | 3 | 2 | | | | |
| Ninth..... | | | | | | | | | | 2 | 10 | 32 | 32 | 29 | 18 | 11 | 3 |

It is seen that not only do first grade children greatly overlap those of the fifth grade, and fifth grade children those of the first year of high school, but that the brightest child in the first grade has all but reached a point in mental ability corresponding to the lowest pupil in the high school. The brightest of the fifth grade pupils are above the median mental level for the first year of high school, and the brightest of the first grade almost reach the median for the fifth grade.

DESIRABILITY OF TESTS IN THE FIRST GRADE

The first grade is the most critical. It is there that retardation scores its worst record, for usually about one-third of the pupils fail of promotion by the end of the first year. Accordingly, it is especially important that in the first grade the raw material with which the school is to work should be correctly evaluated.

By giving Stanford-Binet tests to more than a thousand first grade pupils, and by extensive supplementary data and follow-up work, Dickson¹ has thrown valuable light on the factors which make for the success and failure of beginning pupils. The

¹ See his forthcoming article in one of the early numbers of the JOURNAL entitled "Intelligence tests of first-grade children."

mental ages found ranged from three years to almost eleven years. The part played by mental age is indicated by the correlation, of 0.72 between mental age and the quality of school work as rated by the teacher.² Approximately 90 percent of the pupils below a mental age of six years fail to win promotion. Since approximately one-third were found to be below this level, we seem to have a sufficient explanation for the "piling up" of enrollment in the first grade. On the other hand, Dickson found nearly 15 percent above the mental age of seven and one-half years. Doubtless the large majority of such children could quickly be made ready for the second grade, some of them for the third.

Our only requirement of the entering child has been that of age, which is beside the point. Our real concern is with the mental age. Every child should be tested immediately upon school entrance, and those too immature to stand a reasonable chance of being able to master the work of the first grade should either be sent home or given work of a pre-primary nature. There is no excuse for the postponement of diagnosis until the child has become a confirmed failure.

Dickson's study is not the only one of this kind. In Council Bluffs, Providence, and several other cities, Stanford-Binet tests have been given in large numbers to first grade children. The results appear to be in essential agreement. The child of six and a half to seven years mental age is usually able to master his work; the child of mental age five or five and a half is not.

INEQUALITY OF CLASSES IN THE SAME GRADE

The success of a teacher is judged largely by the absolute standard of work she is able to secure from her pupils. How unfair this may be in individual cases is illustrated by the following facts for five first grade classes tested by Dickson.

² A correlation of 1.00 would mean perfect agreement; a correlation of 0.00 would mean no agreement whatever. A correlation of 0.72 is very high for this type of relationship. It means, of course, that in general the greater the mental age the better the quality of the work.

TABLE VI. INTELLIGENCE OF FIVE FIRST GRADE CLASSES

| Room | Median Mental Age | Median I. Q. | Percent below $5\frac{1}{2}$ (mental) | Percent above 7 (mental) |
|------|----------------------|-----------------|---|--------------------------------|
| A | 6-0 | 87 | 31 | 10 |
| B | 5-7 | 76 | 46 | 5 |
| C | 6-0 | 85 | 20 | 26 |
| D | 7-2 | 108 | 14 | 60 |
| E | 7-8 | 112 | 0 | 71 |

The average mental age of Room E was fully two years above that in Room B, and the median I. Q. 36 points higher. One-third of the pupils in Room A, and half of those in Room B, were incapable of doing standard first grade work. The lack of progress in Room B was so evident that the teacher was in despair and the superintendent doubted her efficiency.

Two fifth grade classes tested by Hubbard yielded a contrast no less striking. The pupils of one class, which we will call class A, ranged in mental age from less than ten years to more than fifteen year; those of class B, from seven and three-fourths years to fourteen years. The median mental age in class A was almost twelve years; that in class B hardly more than ten years. A twelve-year mental age is normal for the sixth grade, a ten-year mental age for the fourth grade. In class A, 44 per cent of the I. Q.'s were 110 or above; in class B only 10 percent. In class A, 19 percent of the I. Q.'s were below 90; in class B, 44 percent. In class A, the median I. Q. was 108; in class B, 91.

The results of the intelligence tests were abundantly confirmed by the corresponding differences between the two classes disclosed by educational tests. Class A was as far ahead of class B in addition, subtraction, multiplication, division, reasoning ability, and spelling as in mental age.

As would be expected, these two classes presented an entirely different picture. The pupils in class A were interested, alert, and above the average in industry; those of class B were inert and unresponsive. The outlook for the two classes is, of course, very different. From 15 to 20 percent of the pupils of class B will never, with any amount of instruction, be able to do the work of the eight grade satisfactorily, and 50 percent are too inferior in endowment ever to complete a four year course in the average

California high school. In class A not far from 80 percent should be able to graduate from high school.

One might suppose that the teachers of these classes would have been keenly aware of the intellectual make-up of their groups. They were not, except in the vaguest sort of way. Each teacher knew that she has some bright and some dull pupils. How bright and how dull was not known. Without an objective standard, such as is furnished by mental tests, the teacher can only judge a pupil by comparing him with the average of his group. ~

RANGE OF INTELLIGENCE IN THE FIRST YEAR OF HIGH SCHOOL

Noting that one-third or more of the pupils who enter high school do not remain to begin a second year of work, Proctor³ decided to attack the problem at its most critical point by investigating the abilities of first year students. He gave Stanford-Binet tests to about 150 pupils and group tests (chiefly Alpha⁴) to 1,349 others. The mental age scores ranged from barely twelve years to nineteen years, approximately 8 percent being below fourteen, and about 23 per cent seventeen years or above. In general, the school marks rose or fell fairly regularly with mental age, although there were individual exceptions due to individual differences in application, health, regularity of attendance, etc. Of the pupils having an average score of B+, not one was below the mental age of fifteen. Two-thirds of the marks earned by pupils having mental ages below fourteen and one-half were C or lower.

From Proctor's study it appears that the standards of work which are maintained in the first year of the average California high school cannot be satisfactorily met by pupils with a Stanford-Binet mental age below thirteen years, and that below the mental age of fourteen the chances of success are not good. It also appears that children with an I. Q. below 80 rarely succeed in entering a California high school, and that those with an I. Q.

³ See a forthcoming article by W.M. Proctor, "Psychological tests as a means of measuring the probable success of high school pupils."

⁴ The group mental test that was given to about 1,300,000 soldiers during the war.

below 90 rarely graduate. A large majority of those who drop out have an I. Q. considerably below 100. The typical high school offers little work which can be mastered by pupils of much less than average intelligence.

Such findings raise the question whether high school standards are too high. Certainly if pupils with inferior ability are to be retained, the high school will have to do one of two things: (1) lower the standard in the present courses; or (2) add other studies which are easier while at the same time educationally worth while. It may be that we have judged the high school too largely by the difficulty pupils encounter in meeting its standards for graduation. Proctor found that 70 percent of those testing below 95 I. Q. failed in more than half of their studies. And more than one-third of all children test below 95. A nation falls short of the true ideals of democracy which refuses to furnish suitable training for a third of its children merely because their endowment does not enable them to complete a course of study which will satisfy the requirements for college entrance. High schools at present are in a sense "class" schools.

MENTAL AGE STANDARDS FOR GRADING

The average age of school entrance in most parts of the United States is not far from $6\frac{1}{2}$ years. Reckoning on this basis, the standard mental age for the different grades would be as follows:

TABLE VII. STANDARD MENTAL AGE

| | | | |
|--------------------|-------------------|-----------------------------------|----------|
| I..... | $6\frac{1}{2}$ to | $7\frac{1}{2}$, or approximately | 7 years. |
| II..... | $7\frac{1}{2}$ " | $8\frac{1}{2}$ " | 8 " |
| III..... | $8\frac{1}{2}$ " | $9\frac{1}{2}$ " | 9 " |
| IV..... | $9\frac{1}{2}$ " | $10\frac{1}{2}$ " | 10 " |
| V..... | $10\frac{1}{2}$ " | $11\frac{1}{2}$ " | 11 " |
| VI..... | $11\frac{1}{2}$ " | $12\frac{1}{2}$ " | 12 " |
| VII..... | $12\frac{1}{2}$ " | $13\frac{1}{2}$ " | 13 " |
| VIII..... | $13\frac{1}{2}$ " | $14\frac{1}{2}$ " | 14 " |
| High School I..... | $14\frac{1}{2}$ " | $15\frac{1}{2}$ " | 15 " |

Children who are located in accordance with these standards are in the large majority of cases found doing work of average quality. If the mental age is much above or below the norms just indicated, the school work is usually correspondingly superior or inferior. This statement is based upon an analysis of

the relation between mental age and school success in the case of 1,936 California children. Of these 1,936, 120 were two or more years above the grade corresponding to their mental age. Only 19 of the 120 were rated as doing anything like satisfactory work. Even in some of these cases the teacher's judgment was open to question. Among the 1,936, there were 234 located in a grade two or more years below the standard for their mental age, and nearly all of these were doing average to very superior work.

In setting forth the above mental age standards for the different grades, it is not implied that all pupils should be graded rigidly according to mental age or that there should be no overlapping of mental ages in adjacent grades. In the first place, the mental age scores yielded by any scale are subject to a certain amount of error. In the second place, the child's school success is determined in part by factors other than intelligence. All these limitations of the mental age method of school grading are freely admitted. It is impossible, however, to justify the present miscellaneous scattering of mental ages over the different grades. If a child is allowed to progress beyond his mental age, the fact should be known and there should be a reason for it. The same rule should hold for the child who is not allowed to advance as rapidly as his mental age would seem to warrant.

For several years Stanford University students have investigated cases in which a marked disagreement was found between mental age and the teacher's estimate of the child's intelligence or school work. These studies have shown that such disagreements are oftener due to the erroneous judgment of the teacher than to any incorrect verdict of the test. In one group of 238 pupils, 34 discrepancies worthy of study were found. The majority of these were not large. In 29 of the 34 cases the quality of the school work as rated by the teacher was poorer than the mental age would seem to warrant. Where discrepancies of this kind occurred they were ordinarily due to physical defects, lack of application, emotional instability, psychopathic heredity, or to such traits as timidity and lack of self-confidence. Exceptional timidity especially has a strong tendency to cause an underestimation of the child's ability. Discrepancies in the

opposite direction were usually due either to exceptional mental application on the part of the child, or to the effect of vivacity, responsiveness, and other favorable personal traits in influencing the teacher's judgment.

One does not need to go far in this line of investigation to convince oneself that objective measures of intelligence are needed to supplement and also on occasion to correct the teacher's judgments regarding the intelligence of her pupils.

INTELLIGENCE TESTS OF SCHOOL LAGGARDS

Notwithstanding the many investigations which have been made of retardation in the last decade, the number of school laggards has decreased but little. Retardation cannot be properly dealt with until its causes are known. The supposed causes usually emphasized are physical defects, irregular attendance, late entrance, poor home conditions, and lack of flexibility in methods of promotion. These "causes" are contradicted by the findings of all who have investigated the subject by the use of mental tests. Of Dickson's first grade pupils who were chronologically eight years old or older, 68 percent were below 80 I. Q. Of Hubbard's fifth grade pupils who were 12 years old or older, 64 percent were below 80 I. Q. Of 174 over-aged children tested in the schools of "X" County, California, 106 were below 80 I. Q. Anyone who desires additional proof need only test a large number of unselected children of a given chronological age, say 11 years, and note the school progress which has been made by those at each mental age. Table VIII is typical.

A table of this kind offers the best possible evidence of the validity of a Binet test as a measure of the child's educability.

INTELLIGENCE TESTS FOR THE GIFTED

With the aid of Mrs. Margaret Hopwood Hubbard, we have gathered extensive data regarding 59 children, most of whom tested above 140 I. Q. The Binet test was taken merely as a point of departure for the investigation.

Most of these gifted children were located in a grade two or three years below that in which they belonged by mental age.

TABLE VIII. GRADE LOCATION OF 263 ELEVEN-YEAR-OLDS BY STANFORD-BINET MENTAL AGE.⁵ CORRELATION IS 0.81

| Mental Age | Grade | | | | | | | | Total |
|------------|-------|----|-----|----|-----|----|-----|------|-------|
| | I | II | III | IV | V | VI | VII | VIII | |
| 18..... | | | | | | | 1 | | 1 |
| 17..... | | | | | | | 3 | 1 | 4 |
| 16..... | | | | | | 1 | 2 | | 3 |
| 15..... | | | | | 2 | 5 | 6 | 1 | 14 |
| 14..... | | | | 1 | 6 | 13 | 2 | | 22 |
| 13..... | | | | 3 | 12 | 18 | 1 | | 34 |
| 12..... | | | 1 | 2 | 22 | 12 | | | 37 |
| 11..... | | | 2 | 10 | 42 | 6 | | | 60 |
| 10..... | | | 6 | 15 | 20 | 1 | | | 42 |
| 9..... | | 2 | 3 | 14 | 6 | | | | 25 |
| 8..... | 1 | 5 | 6 | 2 | | | | | 14 |
| 7..... | 1 | 1 | 3 | | | | | | 5 |
| 6..... | 1 | | | | | | | | 1 |
| 5..... | 1 | | | | | | | | 1 |
| Total..... | 4 | 8 | 21 | 47 | 110 | 56 | 15 | 2 | 263 |

Many of them have since received one or more extra promotions as a result of the test, and all such promotions have been justified by the results. Several were evidently languishing at a level of school work which so far failed to command their best efforts as to affect unfavorably their whole attitude towards school life and school work.

The study further showed: (1) that these children were apparently not below the average in general health; (2) that in the vast majority of cases the intellectual ability was general rather than special or one-sided; (3) that their superiority was also marked in moral and personal traits; (4) that queerness and lack of social adaptability were the exception rather than the rule; and (5) that the superiority in most cases showed itself early in life and had been little influenced by formal instruction. Follow-up tests of the same children strongly suggest that their superiority is likely to be maintained.

We must multiply indefinitely our "opportunity classes" for bright children. Wherever these have been tried they have proved an immediate success. Invariably the children are touched by new life and inspired with new enthusiasm. Gifted

⁵ TERMAN, L. M., *The intelligence of school children*. Boston: Houghton. Mifflin Company, 1919.

children are the nation's greatest asset, one that has not received a thousandth part of the attention it deserves.

THE I. Q. AS A BASIS OF PREDICTION

Mental age is significant because it can be used largely as the basis for school grading. The I. Q. is significant for the reason that it enables us to predict with some degree of approximation what the child's future development will be.

Of course, no one supposes that the score earned in an intelligence test is a perfectly accurate index of an individual's mental status; or that even if we had a perfect measure of intelligence, the ratio between mental age score and age would necessarily remain absolutely constant throughout growth. It is generally recognized that feeble-minded, and particularly epileptic and psychopathic children are subject to erratic changes of I. Q. On the other hand, it is equally true that for the large majority of children, the Stanford-Binet I. Q. shows a marked tendency toward constancy. In the case of 435 I. Q. comparisons of earlier and later tests of the same children the coefficient of correlation was 0.933. Many of these tests were separated by intervals of three to seven years. It was found that the degree of constancy was about the same for bright, average and dull children.

There is nothing else about a child as important as the I. Q. Knowing this, it is possible to forecast with some definiteness what the grade progress will be and what final level of mental development will be attained. The probable error of such predictions should be established.

INTELLIGENCE TESTS IN VOCATIONAL AND EDUCATIONAL GUIDANCE

Vocational guidance usually receives attention only on the eve of the child's departure from school. Thus restricted, it falls far short of its possible value. At every step in the child's progress the school should take account of his vocational possibilities.

We have already seen how frequently the school errs in attempting to force children through courses of study which are beyond their intellectual capacity, and how futile and discouraging such efforts are. It is time that the school should ask not only what it would like to do, but what it can do, for a given pupil. The limits of a child's educability can be fairly accurately predicted in the first school year. By repeated tests these limits can be determined accurately enough for all practical purposes by the end of the child's fifth or sixth school year.

Vocational guidance is not, and may never be, an exact science. Nevertheless, intelligence tests will be of value even if they tell us nothing more than that reasonable success in a given vocation is or is not compatible with the general mental ability which a particular individual possesses. Investigation of the ranges of intelligence in the different types of vocations will make this contribution possible.

Preliminary investigations indicate that an I. Q. below 70 rarely permits anything better than unskilled labor; that the range from 70 to 80 is preëminently that of semi-skilled labor, from 80 to 100 that of the skilled or ordinary clerical worker, from 100 to 110 or 115 that of the semi-professional pursuits; and that above all these are the grades of intelligence which permit one to enter the professions or the larger fields of business. Intelligence tests can tell us whether a child's native ability corresponds approximately to the median for: (1) the professional classes; (2) those in the semi-professional pursuits; (3) ordinary skilled workers; (4) the semi-skilled workers; or (5) unskilled laborers; and this information is of great value in planning a child's education.

SUGGESTIONS FOR THE USE OF MENTAL TESTS

Teachers themselves should be encouraged to use mental tests. All children above the third grade should be given a group mental examination at least every other year. All who score especially high or especially low should be given also an individual examination (Binet, etc.).

All pupils in the first two or three grades ought to be given an individual examination. While the task is large, it is by no

means impossible, and the tests at this time will be more valuable than at any other period of the child's school career. The aim of the school should be to start right with each child. It should not be satisfied to proceed in the dark in the hope that mistakes will be patched up.

It is easier to acquire a certain degree of aptness in giving mental tests than it is to learn how to interpret their results. The teacher or principal cannot give too much study to this part of the problem.

The mental tests should be supplemented by ratings on character traits and by educational tests. Comparison of the various kinds of data regarding a given child will prove of surpassing interest.

In conclusion, it is urged that in the management of a school or a school system the principal and the superintendent will find intelligence tests an indispensable aid in the analysis of their problems, for the simple reason that these problems cannot be dissociated from the quality of material with which the school works.

A NEW KIND OF SCHOOL EXAMINATION

WILLIAM A. McCALL

Teachers' College, Columbia University, New York

There are in the United States about six hundred thousand school teachers. It is a very conservative estimate that each teacher gives on the average twenty examinations a year. This makes twelve million examinations each year. The time required to construct, give, and score each examination will average, say, three hours. This means that thirty-six million hours are spent examining pupils. The data just given will suffice to show the enormous importance of examinations. Without a doubt, they are and will be for some time and may possibly always remain the most important form of educational measurement. Since this is so, it may seem that those of us who are interested in educational measurement have, in our enthusiasm for constructing and standardizing tests, neglected the traditional type of educational measurement. Really, however, this has not been neglect on our part, for standardized tests are nothing but improved examinations. Furthermore we have been learning new technics which will in time react to improve the making of examinations. The purpose of this article is to show teachers how they may make use of one of these new technics of scientific testing not only to improve certain kinds of examinations but also to make examinations a real pleasure instead of an onerous task to both teacher and pupils.

Below is an illustration of a new way to make an examination. It is designed to test a pupil's knowledge of certain facts concerning the physical features of the United States. We have purposely written it hastily in order that it might illustrate certain crudities of construction. Any teacher in the elementary school could do as well and most teachers could do better.

The examination as presented here assumes that the statements whose truth and falsity are to be determined by the pupils have been mimeographed so that a copy of the examination can be placed in the hands of each pupil. The sample examination given below is supposed to have been worked through by a pupil and then scored by a pupil or by the teacher. The *check, cross*

and O mean respectively that the pupil's answer (i. e. his underlining) is correct, incorrect or omitted. Due to limitations of space, only enough of the examination is shown below to illustrate the procedure.

SAMPLE EXAMINATION ON UNITED STATES

Some of the following twenty statements are true and some are false. When the statement is true draw a line under *True*; when it is false draw a line under *False*. Be sure to make a mark for every statement. If you do not know, guess.

1. In general the mountain ranges run east and west True False ✓
2. Most of the rivers flow north True False ✓
3. Mt. Mitchell is the highest point east of the Mississippi River True False X
4. Mt. Washington is higher than Mt. Mitchell True False X
5. The Catskill Mountains are in Maine True False ✓
6. The Cascade Mountains are nearer the Pacific Ocean than the Rocky Mountains.. True False X
7. The Rocky Mountains are nearer the Pacific Ocean than the Appalachian Mountains.. True False ✓
8. The Blue Ridge is in the Rocky Mountains. True False ✓
9. There are more active volcanoes in the west than in the east True False ✓
10. "Old Faithful" is the name of a cyclone which sweeps upward from Texas into Oklahoma. True False X
11. The "Grand Canyon" was cut through the Cumberland Plateau by the Susquehanna River. True False ✓
12. Pike's Peak is in the Rocky Mountains. True False ✓
13. The Mississippi River flows into the Great Lakes. True False ✓

- | | | | |
|--|-------------|--------------|---|
| 14. All the following are tributaries of the Mississippi River: Arkansas, Missouri, Ohio... | <u>True</u> | False | ✓ |
| 15. The Big Sandy is the biggest river in the United States... | <u>True</u> | False | X |
| 16. The Atlantic Ocean is to the east and the Pacific Ocean to the west. | True | False | ○ |
| 17. Canada is to the south and the Gulf of Mexico to the north. | True | <u>False</u> | ✓ |
| 18. The Great Lakes are five in number. | <u>True</u> | False | ✓ |
| 19. It is easier to sink while swimming in the largest lake east than in the largest west of the Mississippi. | <u>True</u> | False | ✓ |
| 20. The central portion of the United States is on the whole more level than the eastern or western portion. | <u>True</u> | False | ✓ |

Number of correct underlinings = 14

Number of incorrect underlinings = 5

Number of omissions = 1

Pupil's score = (number correct) — (number wrong). (A)

Pupil's score = 14—5=9

Let us consider first the reason for expressing a pupil's score as the number correct minus the number wrong. Imagine a pupil who is absolutely innocent of any knowledge of the physical features of the United States. Were such a pupil to take the above test and were he to mark every statement, he would according to the theory of chance mark ten statements correctly and ten incorrectly. The chances of his guessing right or wrong are fifty-fifty or one to one. His score on the above test would be:

$$\text{Score} = 10 - 10 = 0.$$

In short the pupil's knowledge is zero and the method of computing his score gives him zero. Suppose instead that he knows ten statements and guesses at the other ten. Of the ten guessed at he would, according to chance, get five correct and five wrong. That is, even though his real knowledge is ten he will show fifteen

correct ($10 + 5$) and five incorrect. The method of computing his score brings out his real knowledge.

$$\text{Score} = 15 - 5 = 10.$$

A pupil who marks every statement correctly makes a perfect score, viz:

$$\text{Score} = 20 - 0 = 20.$$

Observe that no account is taken of omissions. Only the corrects and incorrects figure in the pupil's score. When the time allowed the pupils to take the test is made short in order to test each pupil's rate of work there will, of course, be many papers showing several omissions each. In all such cases omissions should be ignored, just as we have done above, in computing scores. Even when the time allowed for the test is ample for each pupil to mark every statement, there will still be an occasional instance of omission due to carelessness, or a misunderstanding of instructions or a puritanic conscience against increasing the score by gamble guess-work even when the instructions urge guessing. When the time is ample for even the slowest pupils and when all are instructed to mark every statement it is much more convenient to compute a pupil's score according to the following formula:

Score = (number of statements)—2 (number marked incorrectly) (B).

To derive this formula from the formula marked (A) above, note that since there are no omissions, the number of corrects (C) plus the incorrects (W) equals the number of statements (N). Hence we may write $C=N-W$. Substituting this in formula (A) gives:

$$\text{Score} = (N-W)-(W)=N-2W.$$

If there are 20 statements in the test and if five are marked incorrectly,

$$\text{Score} = (20) - 2 (5) = 10$$

Formula (A) gives the same results, i. e.:

$$\text{Score} = 15 - 5 = 10$$

Both formulae give identical results provided there are no omissions. Formula (A) is basic and should be used when there are omissions. Formula (B) should be preferred when there are no omissions or when they are present only in negligible amount. Formula (B) is much more convenient. The

first number is always the same and since the second number is twice the total statements marked incorrectly, it is only necessary to score and total the errors.

It is very difficult for some people to believe that such a test as we have outlined above does anything more than give the highest score to the luckiest guesser. They look with an eye of suspicion upon this thing we call *chance*. Being in a position which offered excellent opportunity, namely treasurer of a Sunday school, we once tossed pennies for heads or tails fifty thousand times. The results came out 25,000 heads and 24,999 tails. Had there not been a miscount somewhere the two would doubtless have come out exactly even. We had occasion to *watch* two summer-school teachers engage in that soul-absorbing, nerve-racking game of chance called *matching pennies*. Each began the summer with a special bag of one hundred pennies. They matched for several minutes daily. The last we heard they were still matching pennies and chance had prevented either from getting complete possession of the other's one hundred pennies. Chance is fatally exact when the pennies or the statements in the test are numerous. The opportunities for injustice in scores multiply in proportion as the number of statements is reduced. Hence there should be as many statements in the test as practical limitations will permit.

The possibility of unreliability of pupil scores is one which this sort of examination shares with all other sorts of tests and examinations. Last summer we gave a test of this kind to our class in educational measurement. As luck would have it two visitors were present and were persuaded to take the examination along with the rest. There were forty statements in the test. The score of one visitor was zero and of the other - 2 (when scores are negative they may be considered zero. It is probably the accidents of chance which show an individual to know less than nothing, though a genuine negative score is possible where an individual has been taught errors.) The highest score made by regular members of the class was 36, the lowest score was - 4, and the next to the lowest score was 6. The student who made - 4 admitted that he was not sure of the truth or falsity of a single statement and that unexpected outside demands had until then

prevented him from studying. Even though he was not given an F at the end of the summer session and even though the course is not required, he appeared at the college this last September and registered for the course again. The student who made a score of 6 registered late. He had been in the course only four or five days. The individual who made the highest score was George Melcher, Director of the Bureau of Research, Kansas City, Mo. All were in agreement that he knew more about educational measurement than anyone else in the class.

During a semester course in educational psychology we gave four examinations. The first was of the *True-False* variety, the second was of the traditional *What* and *Discuss* kind, the third was *True-False*; and the fourth was traditional like the second. Each test lasted about thirty minutes. The half-dozen best and poorest students according to the first test remained with negligible variation the half-dozen best and poorest in all the examinations no matter what the type. An amusing and pathetic incident of the course was the challenge of the poorest students to the rest of the class. Those who had been lowest all through the term proposed in the last test to displace some of those who had been consistently above them. There was a sudden rise in the quality of everyone's work, but the lowest remained the lowest. These words should not be interpreted as meaning that the *True-False* test is unusually reliable. So far as our inadequate data go, there is no reason to suppose that it is much more reliable than the traditional examination.

There are a few suggestions which will help teachers who may wish to use the *True-False* test. First, the teacher should so construct the test that it will contain approximately the same number of true and false statements. A clever pupil may get a higher score than he deserves if he discovers there are many more true statements than false statements in the test or *vice versa*. Suppose there are many more true statements than false statements and suppose some pupil discovers this by observing the statements that he knows, or by observing the teacher's bias for writing true statements instead of false ones. Naturally when he does not know what to mark he will mark *True* thereby securing a little larger score than his ability justifies. It is probably

by just such utilization of the errors of others that the intelligent get through life so much more smoothly than the stupid. On the other hand, the teacher should not have exactly the same number of true and false statements each time, because this will invite clever pupils to count back to see how many more true statements have been marked than false statements. Sometimes there should be more true statements, sometimes more false statements, sometimes the same number of each. A regular plan either as to content or arrangement should be carefully avoided. All the true statements should not come first, neither should the true and false statements be alternated as a regular plan. Let chance determine how many shall be true and how many shall be false and in what order the true and the false shall come.

Second, the teacher should be careful to keep out of the test all ambiguous statements. Statement number 18 in our sample test is somewhat ambiguous. It says: "The great lakes are five in number." Since *great lakes* is not capitalized a pupil might very legitimately interpret this to include Great Salt Lake and others. It will later be difficult to satisfy this pupil that his score should suffer because of the construction he gave this sentence. If the teacher will study her mistakes in this respect she will soon learn how to reduce such ambiguities. As any teacher can testify, the danger of ambiguities of wording are not peculiar to this test. This type of test does not, however, give a pupil an opportunity to reveal just what interpretation he places upon each statement. If the teacher follows the procedure of having pupils score their own or each other's paper all cases of serious ambiguity will be discovered. Statements which are particularly flagrant in this respect can be omitted in scoring.

Third, the teacher should inspect not only this but any sort of test from the point of view of just what the test measures. Statement 19 in our sample test illustrates our point. Our purpose is to test whether the pupil knows that the largest lake west of the Mississippi River contains more salt than the largest lake east of the Mississippi. Instead of measuring this we may be testing whether a pupil knows that it is easier to sink in fresh water than in salt water. Complex wording, unfamiliar terms, the use of

negatives, all tend to make the test a linguistic one. Simple, brief statements without negatives are best.

Fourth, the teacher may so construct the examination as to force pupils to guess wrong due to the power of suggestion. This probably explains why statement 15 was marked wrongly. The pupil doubtless argued to himself that since the river is named the Big Sandy it probably is the biggest river in the United States. The influence of having many suggestive statements in the test is to make the examination more difficult. It operates to give to the pupil who knows nothing at all in the test a large negative score instead of a zero score and it penalizes rather heavily the pupil who does much guessing, for every time he allows himself to be suggested in the wrong direction a point is subtracted from the score he has already made by what knowledge he has. In other words, the suggestive statements make the gap between those who know much and those who know little wider than it otherwise would be. Whether a pupil should be specially penalized for allowing himself to be maneuvered wrong is an arguable question. There may be situations where it is eminently desirable to determine whether pupils know what they know so well as to be able to resist suggestion. But we believe that, in general, it is best to avoid suggestive statements. The ideal should be so to construct the examination that any pupil who knows absolutely nothing about the test will make a score of zero.

So much for the construction of the examination. How shall it be applied? The best way, as shown by our sample, is to print, mimeograph, or otherwise duplicate, the examination, and place a copy in the hands of each pupil. But there are numerous schools which lack duplicating machines. For teachers in these schools some other means for applying the test must be found. Any one of the following methods may be used. First, the entire test may be copied word for word by the pupils and then marked. This is tedious and time consuming. Second, the entire test may be written on the blackboard by the teacher. Each pupil could number a blank page of paper to correspond to the numbered statements, and then write *True* or *False* after the appropriate numbers. The only objection to this suggestion is the inconvenience of writing all the statements on the blackboard. Third,

the pupils may be asked to copy on blank paper, 1, 2, 3 and so on, according to the number of statements. The teacher can then read orally statement number 1 and instruct the pupils to make a check after the number 1 on their paper if the statement is true, but to make a cross if the statement is false. This is easily the most convenient way to give the examination. The chief objection to this last method is the difficulty some pupils have in apprehending statements presented orally, particularly if they are long and complicated. When the statement is presented visually the pupil has an opportunity to go back to it enough times to realize his possibility of understanding it. By one or another of these methods it is possible for any teacher anywhere to make use of this type of examination.

How should the *True-False* examination be scored? If a copy of the test has been placed in the hands of each pupil, the teacher can take an unused test sheet, fill it out correctly, lay the correct column of answers beside the pupil's column of answers, and quickly mark whether the pupil's answers are correct or incorrect. If a copy of the test has not been placed in the hands of each pupil, but each has instead written *True* or *False*, or made a check or cross after the number of each statement, the teacher can take a page of paper similar to that on which each pupil has indicated his answers, copy the numbers just as they are and just as they are spaced on each pupil's paper, write after each number the correct answer to the statement of that number, place this column of correct answers beside the column of pupil answers and mark those which are correct and incorrect. This last scoring method presupposes that pupils have used ruled paper, and that each has written his numbers in a vertical column according to a particular spacing recommended by the teacher. Last and best each pupil can score his own or his neighbor's paper. It is better for him to score his own.

If the method of pupil scoring is adopted, the teacher should read the correct answers while the pupil checks his own. If the pupil does not have a copy of the statements before him, the teacher should read each statement before giving the correct answer, in order that the pupil may know what statements he got correct or incorrect. When all the pupils' answers have been

marked and when all their scores have been computed and recorded on their examination paper, the teacher should ask all pupils who missed statement number 1 to hold up their hands, and then all pupils who missed number 2 to hold up their hands and so on. The teacher should make a record of the number of pupils missing each statement, and then collect all papers.

But why should this sort of examination be given at all? Wherein is it superior to the examination method in common use? In the first place, the *True-False* examination permits a teacher to cover a wider field of subject matter or a wider range of ability per unit of time. It may be made more representative of the total field of the pupils' study. In the case of the traditional examination the teacher is forced to select a very small number of questions. When we were students, almost as much of our ingenuity went into divining the kind of questions the teacher would ask as into reviewing. Now that we are teachers we have no reason to suppose that this practice has ceased. Rivalry among pupils in taking the traditional examination is largely a contest in guessing. The *True-False* examination can ask so many questions in a brief time that the only hope of the pupil is to study everything everyday and review everything when he reviews. Further, this extensive examining is fairer to pupils. The pupil has some basis for the time-honored excuse that he was so unlucky as to know everything except the particular questions asked by the teacher. The new type of examination rules this alibi out of court.

The use of this type of examination is likely to improve the relation between teacher and pupils. The traditional examination endangers a pleasant relationship because pupils more or less justly suspect that the score they make depends almost as much upon their conduct as upon their product. According to the pupils' suspicions the teacher uses her control over scores to vent a personal spite against the bad boy, and to express her appreciation of the "pet."

There is no question but what the pupils are partly correct. The traditional examination is so subjective that it is well-nigh impossible for the teacher to prevent her judgment from being influenced by fluctuations in her disposition or from being colored

by her affection for the pupil whose paper she is scoring. The new examination does away with these prolific sources of unpleasantness. It is so objective that the pupil can see for himself just how well he did. It permits self-scoring, and it convinces a pupil that the score he gets is the score he deserves.

The *True-False* examination is more enjoyable for the pupils. "Children cry for it" may be a bit exaggerated, but at any rate they hate it less. It offers an opportunity for a contest where the rules are fair, a chance for a larger degree of participation in the examination. It is agonizing for a pupil to describe at great length a knowledge which he does not possess in the hope that his command of English will camouflage his lack of information. Here is a question which was asked in a recent examination in educational measurement:

Which three of the tests described by Whipple do you think would be of most service in an elementary school, if your school had a psychologist to apply them?

Consider the perspiration it must have cost a student to perpetrate this answer:

The tests described by Whipple embraced most of the difficulties that would be embraced in problems of classroom instruction. I think his tests embrace a great variety of methods of approach and it seems difficult for me to think of just three to whom the presence of a psychologist in a school would give help. I would think it would be the tests in which knowledge of the workings of a child's mind and its growth and development would be most apparent since those not particularly trained might focus on others not of this kind. I fear it would be unwise to specifically mention just three when the number is so great which would fulfill all these requirements. Every teacher to be a psychologist would help all classroom measurement work of whatever kind greatly, I know since we cannot know of the influence of a test upon any group except by the mental reaction produced.

The *True-False* examination is also more enjoyable for the teacher because the scoring is easy, rapid and automatic when she does the scoring, and far more rapid when the pupils do the scoring. The pupils cannot well assist in scoring the traditional examination, and for the teacher to score forty verbose examination papers is time-consuming drudgery. Every moment of the time while scoring, the teacher must be profoundly concentrating upon what she is reading, for much of the time she must be sepa-

rating the chaff from the wheat where the chaff is cleverly painted to look like wheat. And along with this is a continual emotional strain caused by her resistance to the temptation to underscore some and overscore others.

The *True-False* examination is more educative to the pupils. The proposition that pupil scoring will relieve the teacher of much obnoxious drudgery, does not justify the inference frequently made that what is non-educative drudgery for the teacher will also be noneducative drudgery for the pupils. On the contrary we are of the opinion that the most favorable teaching opportunity that ever comes to a teacher is the period immediately following an examination. The pupil's interest to know what parts of the examination he missed and what he got correct is then at white heat. Witness the interested discussion among pupils immediately following an examination. It is inexcusable neglect of an educational opportunity not to capitalize these precious moments for correcting erroneous ideas, clinching right ideas, and filling up mental spaces where ideas are not. These values can best be realized by having each pupil score his own paper and by stopping to discuss points where pupils have trouble. Of course not every correct answer indicates knowledge, but the pupil himself usually knows when he knows. This examination is also more educative, because it is likely to be given more frequently. The experience of Kirby, Courtis, and others with practice tests shows that a pupil learns more during testing periods than during teaching periods. We really teach when we test. This examination covering as it can a wide range is an ideal method of review. It reveals to the pupils just where their difficulties lie. Testing is one of the best ways of teaching.

The *True-False* examination gives the teacher a fuller knowledge of conditions. The educative value of testing is so great that testing should be much more frequent than is now the case. If a method of testing is available which involves no drudgery to anyone, testing is likely to become more frequent, and this means more complete and timely information about the abilities and difficulties of the various pupils, and about the successes and failures of teaching efforts. We have already suggested that the teacher keep a record of the number or percent of pupils

missing each statement in the examination. This record will show what things have been well-learned or poorly learned, and well taught or poorly taught. Also it is a good thing for a teacher to check the effectiveness of her teaching. This can be done by finding the average of all pupils' scores and by comparing this average with the total number of statements in the examination or at least the total number of facts the teacher had really attempted to teach the pupils. If the average score is 20 out of a possible 40, the teaching efficiency is 50 percent.

Finally the *True-False* examination is a genuine honesty test, and shows the beginnings of a technic for measuring in satisfactory fashion this valuable character trait. Occasional and unannounced rescoring of each pupil's paper by his neighbor will catch the persistent cheat. It is better that he be discovered in school than in court. His discipline can usually be left to his fellow pupils, over whom he was attempting to gain an advantage dishonestly.

We have listed what appear to be the chief advantages of this type of examination or any other examination which is similarly objective. Most of these claims rest upon logical probability and a limited experience and not upon experimental data. This last is needed and will follow in time.

There are some limitations which have not yet been discussed. It is claimed first that this examination does not require the pupil to demonstrate a power to organize his materials. This is true in the sense that the pupil does not *describe in writing* a complicated mental organization but a statement can be so worded as to require an exceedingly complex mental organization before a correct answer can be unfailingly given. Consider the mental organization that must precede a correct answer to this simple statement: "If the trade winds blew east Peru would have luxuriant flora." If it is desired to test a pupil's power to word his thought a composition test may be given.

Again, it is claimed that this examination can test knowledge but not skill, knowledge but not the ability to do. Even skills, however, can be tested by this type of examination. To reason that trade winds blowing east would be warm, would absorb moisture from the Pacific, would become chilled in passing over

the Andes, would consequently deposit a heavy rainfall for Peru, which taken in conjunction with the equatorial climate would produce a luxuriant flora, is one sort of skill which this examination will test. Mathematical skills and the like which are too complicated to describe may be tested in at least two ways. An example or problem may be stated together with an answer. The pupil's task will be to determine by working the problem whether the answer given is true or false. Or instead, the teacher may work the problem on the blackboard for all the pupils and have them indicate whether her process was correct or incorrect.

Finally it is claimed that the teacher needs to know why a pupil is unable to answer the question about Peru and its flora. The *True-False* examination does not show just where the pupil's reasoning process went wrong or stopped altogether. It is not diagnostic. This criticism has some force. An examination should be as diagnostic as possible. If a teacher wished to know where the pupil's process broke down she could give a subsequent more detailed examination of this type. The statement: "The trade winds are warm winds" or "Warm winds have a larger capacity for water than cool winds," etc., would reveal whether the pupils were acquainted with the basic principles, facts and the like necessary to reason out the correct answer to: "If the trade winds blew east Peru would have luxuriant flora."

The traditional examination has certain advantages which will doubtless continue its existence. The *True-False* examination is a herald of newer and better types of examinations. But even now we have in the *True-False* examination one that may be used by any teacher anywhere to great advantage.

HURDLES, A SERIES OF CALIBRATED OBJECTIVE TESTS IN FIRST YEAR ALGEBRA

MURRAY A. DALMAN

Director Department of Reference and Research, Indianapolis, Indiana

The purpose of this report is to show how the teachers of algebra in one high school solved one of their problems. It is because the material is in daily use and the circumstances common that the discussion may be of value to teachers searching for help in one of their many difficulties. The method is in continuous use in the Emmerich Manual Training High School of Indianapolis, Indiana—a school in which things without value are eliminated.

In common with most mathematics teachers, the teachers of first-year algebra in that school were concerned with the great number of failures. Because they were excellent teachers, well-trained, professionally interested, alive to every problem of the department, they were determined to effect a remedy if possible. They tried several remedial plans which either did not succeed, or were given up on account of administrative difficulties.

1. Classes were made up on the basis of ability in algebra, and pupils were redistributed during the term on this basis; but the crowded condition of the school made this division difficult of administration, and the plan was itself of doubtful value.

2. Classes were made up of the pupils who were failing. In some of these sections, special emphasis was placed upon the extension of arithmetic into the generalities of algebra. In others an essential minimum of algebra was taught; but the minimum varied with different teachers, and the results were widely different.

3. Student assistants from the upper classes were appointed to coach the dull and failing and made a brave attempt to bring up the laggards, but the results were about the same.

4. Supervised study classes were instituted and the brighter pupils did better. But the number of failures was nearly as great as before, and the doubled cost could not be justified. The administrative difficulty on account of crowded conditions again interfered.

5. A series of mental tests was given and, by using those tests which showed a sufficient correlation with teachers' ratings in algebra, it was found possible to foretell with reasonable accuracy a pupil's ability and probable degree of success in algebra. A set of "diagnosis data" was worked out so that the mental test scores could easily be read in terms of algebra grades. This diagnosis was made at the beginning of the term in an effort to bring each pupil to function to the limit of his ability in algebra, or to find a reason for not being able to do so. This brought considerable improvement in the condition; but it was a rather cumbersome method of procedure. The results of the tests were good; but the testing of five to seven hundred 9B pupils, the scoring of papers, and the diagnosis of each case was quite a task, and the diagnosis could never be absolute. There were many exceptions.

These attempts to bring each child to work worthy of his ability led to a discussion of the course of study in algebra and the requirements for passing. The needs of the department were brought out as follows:

1. A course consisting of "minimum essentials," with a uniform requirement for passing.
2. More difficult work and courses for pupils of greater ability.
3. A statement of requirements for the pupils and teachers.

In revising the course of study and its administration, there were four facts to be taken into consideration:

1. A text adopted for the state.
2. The standard marking system.
3. The report card periods.
4. The administrative problems and difficulties of a crowded school.

The "minimum essentials" course was intended to be such a course that all pupils who completed it would have a complete and comprehensive view of the principles of first year algebra, although they would be unable to manipulate quantities skilfully enough to work the difficult exercises.

This course was reached by an analysis of the ends to be attained in a year of algebra. It was a matter of determining,

from a consideration of the end of the course, what knowledge of principles and what skills in manipulation were absolutely necessary for successful or barely passing work. The selection of principles was a comparatively simple matter, but the second step of determining the skills in manipulation required was not so simple. It was not easy, for example, to be sure that there were no especially difficult points in the requirements—no points, that is, under any topic which could not be handled by pupils who had been successful in the previous topics.

Then there remained the difficulty of arranging the material so that both pupils and teachers would know definitely what the minimal requirements were. To solve this problem and also to provide successive standards for pupils of successively greater abilities, a series of objective tests was planned.

The standard marking scheme of the school was A+, excellent, 95 to 100 percent; A, very good, 90 to 95 percent; B, good, 80 to 90 percent; C, fair or passing, 70 to 80 percent; and D, failure, below 70 percent. Tests were accordingly made out to fit the standards, C, B, A, and A+. A chart was made (see Figure 1) for posting in each classroom so that every pupil who passed a test could be publicly given credit for it. Successes were indicated by heavy dots placed in certain positions. Such an entry in column R, showed that the work to date had been passed. No recognition was given a pupil in this column if he had failed in the work of a former period and still had that work to make up.

In the succeeding columns were entered for each child the grade value (C, B, A, or A+) of the test passed and the number of attempts necessary to pass it. The chart provides for four attempts under each test. If more than four are required the number of attempts may be entered as in the case of the first pupil designated as A. S.

The chart reads as follows: I. C. made a grade of A, or 90, passing each test on the first trial; L. H. made a grade of B, but only after three attempts at the C test and four at the B; etc. The record for this class shows that 2 were failed, 7 received a grade of C, 8 of B, 5 of A, and 2 of A+.

To the regular form of chart has been added in this case, column M, for the diagnosis of algebraic ability made at the

FIGURE 1. CLASS RECORD CHART, SHOWING PREVIOUS DIAGNOSIS (M) RECORDS ON FORMER TOPICS (R) AND TESTS COMPLETED (C, B, A, A+). ALGEBRA II, 5TH HOUR, ROOM 12, M. D., TEACHER

| | M | R | C | | | | B | | | | A | | | | A+ | | | |
|------|----|---|---|---|---|---|---|---|---|---|---|---|---|---|----|---|---|---|
| | | | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 |
| I.C. | 60 | ● | ● | | | | ● | | | | ● | | | | | | | |
| L.H. | 79 | ● | | | ● | | | | | ● | | | | | | | | |
| D.L. | | ● | | ● | | | ● | | | | | | | | | | | |
| F.L. | 82 | ● | | | ● | | | | | | | | | | | | | |
| S.M. | 84 | ● | ● | | | | | ● | | | | | | | | | | |
| M.R. | | ● | ● | | | | ● | | | | | ● | | | | ● | | |
| B.R. | 80 | ● | ● | | | | | ● | | | | | | | | | | |
| M.R. | 76 | ● | | | ● | | | | | | | | | | | | | |
| B.R. | 98 | ● | ● | | | | ● | | | | ● | | | | | ● | | |
| A.S. | 76 | ● | | | | ● | | | | | | | | | | | | |
| M.S. | 92 | ● | ● | | | | ● | | | | ● | | | | | | | |
| R.S. | 72 | ● | | | ● | | | | | ● | | | | | | | | |
| A.S. | | ● | | | ● | | | | | ● | | | | | | | | |
| E.S. | 72 | | | | | | | | | | | | | | | | | |
| J.S. | 76 | ● | | | | | | | | | | | | | | | | |
| D.S. | 84 | ● | | | ● | | | ● | | | | | | | | | | |
| R.S. | 79 | ● | | | ● | | ● | | | | | | | | | | | |
| A.S. | 89 | ● | ● | | | | ● | | | | ● | | | | | | | |
| R.T. | | ● | ● | | | | | | | | | | | | | | | |
| H.T. | | ● | | ● | | | | | | | | | | | | | | |
| N.V. | 78 | ● | | | ● | | | | | | | | | | | | | |
| L.W. | 88 | ● | | | ● | | | | ● | | | | ● | | | | | |
| R.W. | 80 | ● | | | ● | | | | | | | | | | | | | |
| R.W. | 81 | ● | ● | | | | ● | | | | ● | | | | | | | |

beginning of the term before. Only those pupils who were in attendance at that time have records.

The only instance of complete disagreement between the diagnosis and the grade is in the case of I. C. For this reason her case was investigated. It was found that her parents were deaf and dumb, and that she had been kept closely at home. She had entered school late, and had never had a normal mental growth. She failed in all other subjects because she devoted all her time to her algebra.

The ninth pupil (B. R.) with a diagnosis grade of 98 reached a grade of A+ on the first attempt at each set of tests.

Figure 2 is a comparative arrangement of sample tests for each topic into which the course in first year algebra was divided. The tests are given in the vertical columns in the order in which the topics are taken up, and in the horizontal columns in the order of difficulty in the tests in each topic. By this table it is possible to see the minimal requirements for passing, the relatively greater demands for a B, A, or A+ grade as a whole or to compare the relative requirements for each grade in any topic.

In these tests the passing grade is 100 percent and the exercises are calibrated with this in view. This avoids questions of credit for partial solutions, and differences of judgment on the part of teachers, and also makes sure that each pupil who passes a test is familiar with all the principles involved and can accomplish the manipulations necessary for success in the next topic.

There was no special endeavor to make all the exercises of any given test of exactly equivalent difficulties (although this, of course, is approximately true), but rather to make each test cover all the principles and conditions underlying the particular topic involved. As between tests of the same grade (C, B, A, or A+) the effort was to make them equivalent to each other, in proportion to the increasing knowledge of algebra.

There was no attempt to make the test for pupils of greater ability cover all the points and principles of the C tests although this is true in the main, but rather to make them cover the additional information and skill in manipulation gained by the pupils of greater ability.

After this preliminary work had been completed, the result was tried out in two classes. This trial showed that the analysis of difficulty in making the tests was faulty and that the minimum set was higher than a proportion of the pupils could reach. This was due to the fact that under the traditional scheme of grading the teachers had no proper conception of the absolute knowledge of the pupils, and that the imaginary standards were never reached. In other words measuring on a basis of the accomplishment of approximately three fourths or 70 percent of material of a wide range of difficulty, had led teachers to believe that,

FIGURE II. THE TESTS, SHOWING THE REQUIREMENTS FOR PASSING AND FOR HIGHER GRADES IN FIRST YEAR ALGEBRA—
CLASSIFIED BY TOPICS AND DIFFICULTY

FIRST SEMESTER—ALGEBRA I

| TOPIC | C | B | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|----------------------|--|---|------------------|--------------|---------------------------------|--------------------|------------------|------------------|--------------|--|----------------|----------------------------|-------------------|------------------|----------------------|--------------|--|----------------|------|---|-----|-----|----|---|----|---|----|---|----|----|----|----|----|---|-----|----|---|----|----|----|----|----|
| Addition | <p>Add:</p> <table><tr><td>(1)</td><td>(2)</td><td>(3)</td><td>(4)</td><td>(5)</td></tr><tr><td>-2</td><td>-6</td><td>7</td><td>3</td><td>0</td></tr><tr><td>-7</td><td>9</td><td>-11</td><td>-3</td><td>5</td></tr></table> <p>6. $-2+7-4=$ 7. $-9+5+1=$ 8. $7-4-3=$</p> | (1) | (2) | (3) | (4) | (5) | -2 | -6 | 7 | 3 | 0 | -7 | 9 | -11 | -3 | 5 | <p>Add:</p> <table><tr><td>(1)</td><td>(2)</td><td>(3)</td><td>(4)</td><td>(5)</td></tr><tr><td>-5</td><td>5</td><td>-2</td><td>6</td><td>-8</td></tr><tr><td>4</td><td>-5</td><td>-7</td><td>-3</td><td>-7</td></tr><tr><td>-3</td><td>4</td><td>-10</td><td>-7</td><td>6</td></tr><tr><td>-1</td><td>-3</td><td>+7</td><td>+8</td><td>-1</td></tr></table> <p>6. $-4-3+5-7=$ 7. $-6+5-7-2=$ 8. $3-9+5-3=$</p> | (1) | (2) | (3) | (4) | (5) | -5 | 5 | -2 | 6 | -8 | 4 | -5 | -7 | -3 | -7 | -3 | 4 | -10 | -7 | 6 | -1 | -3 | +7 | +8 | -1 |
| (1) | (2) | (3) | (4) | (5) | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| -2 | -6 | 7 | 3 | 0 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| -7 | 9 | -11 | -3 | 5 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| (1) | (2) | (3) | (4) | (5) | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| -5 | 5 | -2 | 6 | -8 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 4 | -5 | -7 | -3 | -7 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| -3 | 4 | -10 | -7 | 6 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| -1 | -3 | +7 | +8 | -1 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Evaluation | <p>If $a=2$, $b=-3$, $c=1$, $x=-5$, $y=0$ and $z=6$, find value of:</p> <table><tr><td>1. $a+ab-c$</td><td>5. x^3-5a^2</td></tr><tr><td>2. $-2abx$</td><td>6. $\frac{az}{3}+\frac{a^2}{4}$</td></tr><tr><td>3. $2b^2$</td><td></td></tr><tr><td>4. $5ayb$</td><td></td></tr></table> | 1. $a+ab-c$ | 5. x^3-5a^2 | 2. $-2abx$ | 6. $\frac{az}{3}+\frac{a^2}{4}$ | 3. $2b^2$ | | 4. $5ayb$ | | <p>Grade of B given to all children who pass C test on third attempt.</p> | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 1. $a+ab-c$ | 5. x^3-5a^2 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 2. $-2abx$ | 6. $\frac{az}{3}+\frac{a^2}{4}$ | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 3. $2b^2$ | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 4. $5ayb$ | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Subtraction | <p>Subtract:</p> <table><tr><td>(1)</td><td>(2)</td><td>(3)</td><td>(4)</td><td>(5)</td><td>(6)</td></tr><tr><td>-3</td><td>-9xy</td><td>8</td><td>0</td><td>-ac</td><td>2mn</td></tr><tr><td>4</td><td>-2xy</td><td>-7</td><td>-7bc</td><td>ac</td><td>10mn</td></tr></table> <p>(7) $\begin{array}{r} a-3b+c \\ 3a-b+5c \\ \hline \end{array}$ (8) $\begin{array}{r} 2k \\ -k+5m \\ \hline \end{array} +3r$</p> <p>9 From $3x+2y-4z$ take $z+y-2x$ 10 Subtract $3m-r+2p$ from $m+r-5p$</p> | (1) | (2) | (3) | (4) | (5) | (6) | -3 | -9xy | 8 | 0 | -ac | 2mn | 4 | -2xy | -7 | -7bc | ac | 10mn | <p>1. Subtract $2x+x^2-5-x^2$ from $1-3x^2+2x-x^2$ 2. From $\frac{1}{2}a+\frac{1}{3}b-c$ take $\frac{1}{4}a-\frac{1}{5}b+\frac{1}{6}c$ 3. From $2a$ subtract $b+c$ 4. Subtract $x+2y$ from $3x-z$</p> | | | | | | | | | | | | | | | | | | | | | | |
| (1) | (2) | (3) | (4) | (5) | (6) | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| -3 | -9xy | 8 | 0 | -ac | 2mn | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 4 | -2xy | -7 | -7bc | ac | 10mn | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Equations | <p>1. The perimeter of a rectangle is 160 yards. The base is four times the altitude. Find the dimensions of the rectangle. 2. The sum of two numbers is 78. The greater exceeds the less by 4. Find the numbers. 3. Solve $2x-(3x+6)=11-2x$ and prove the result. 4. Solve $5(x-2)+5=20$ and prove the result.</p> | <p>1. A farmer wishes to enclose a rectangular field for a pasture, making it 24 rods wide. He wants to make it as long as possible, using 168 rods of wire fencing which he has on hand. How long can he make it? 2. The sum of two consecutive integers is 69. Find them. 3. The sum of the ages of A and B is 70 years. In 5 years A will be three times as old as B. Find their ages. 4. Solve $\frac{x}{2}-\frac{x}{5}=3$. Prove.</p> | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Division | <p>Divide:</p> <p>1. $x^2-2x-15$ by $x-5$ 2. $6x^2+5x+6$ by $3x-2$ 3. a^3+b^3 by $a+b$</p> | <p>Divide:</p> <p>1. $6a+5-23a^2+12a^3$ by $4a-5$ 2. $81m^4-1$ by $3m+1$</p> | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Factoring | <p>Find prime factors of:</p> <table><tr><td>1. $3x^2+15x+18$</td><td>5. $am+bm+ar+br$</td></tr><tr><td>2. m^4-a^4</td><td>6. $1-b^2$</td></tr><tr><td>3. $3c^2+6cy+3y^2$</td><td>7. $x^5+x^4+x^3$</td></tr><tr><td>4. $6a^2+22a+12$</td><td>8. b^2+c^3</td></tr></table> | 1. $3x^2+15x+18$ | 5. $am+bm+ar+br$ | 2. m^4-a^4 | 6. $1-b^2$ | 3. $3c^2+6cy+3y^2$ | 7. $x^5+x^4+x^3$ | 4. $6a^2+22a+12$ | 8. b^2+c^3 | <p>Find prime factors of:</p> <table><tr><td>1. $16k^4-b^4$</td><td>5. $2a^2b^3-2a^2b^2+2a^2b$</td></tr><tr><td>2. $48a^2+72a+27$</td><td>6. $bx+by-mx-my$</td></tr><tr><td>3. $100x^2-310x+150$</td><td>7. $1-27m^3$</td></tr><tr><td>4. $3x^2-18x+24$</td><td>8. $64x^3+a^3$</td></tr></table> | 1. $16k^4-b^4$ | 5. $2a^2b^3-2a^2b^2+2a^2b$ | 2. $48a^2+72a+27$ | 6. $bx+by-mx-my$ | 3. $100x^2-310x+150$ | 7. $1-27m^3$ | 4. $3x^2-18x+24$ | 8. $64x^3+a^3$ | | | | | | | | | | | | | | | | | | | | | | | | |
| 1. $3x^2+15x+18$ | 5. $am+bm+ar+br$ | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 2. m^4-a^4 | 6. $1-b^2$ | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 3. $3c^2+6cy+3y^2$ | 7. $x^5+x^4+x^3$ | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 4. $6a^2+22a+12$ | 8. b^2+c^3 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 1. $16k^4-b^4$ | 5. $2a^2b^3-2a^2b^2+2a^2b$ | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 2. $48a^2+72a+27$ | 6. $bx+by-mx-my$ | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 3. $100x^2-310x+150$ | 7. $1-27m^3$ | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 4. $3x^2-18x+24$ | 8. $64x^3+a^3$ | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

FIGURE 2—Continued

FIRST SEMESTER—ALGEBRA I

| A | | | | | A+ | | | | TOPIC |
|--|----------------------|-----|-----|-----|--|---------------------------|-----|-----|-------------|
| Add: | | | | | Add: | | | | Addition |
| (1) | (2) | (3) | (4) | (5) | (1) | (2) | (3) | (4) | |
| - 3 | 17 | -14 | 12 | -15 | 17 | 36 | -12 | 32 | |
| + 9 | -12 | 13 | -10 | 13 | 23 | -18 | -28 | -41 | |
| - 8 | -13 | 20 | 7 | 19 | -12 | 45 | 17 | 13 | |
| +12 | 15 | -18 | - 6 | -12 | 42 | -16 | -19 | -12 | |
| -15 | -20 | -12 | -17 | 10 | -18 | -29 | +22 | -48 | |
| | | | | | 30 | | | | |
| | | | | | - - | | | | |
| 6. $19+16-20+12-10=$ | | | | | 5. $-35-12+19+15-27=$ | | | | |
| 7. $20-8-12+16-5=$ | | | | | 6. $15+19-13+42-34=$ | | | | |
| 8. $19-7+8-4+3=$ | | | | | | | | | |
| Grade of A given to all children who pass C test on second attempt. | | | | | Grade of A+ given to all children who pass C test on first attempt. | | | | Evaluation |
| 1. From $2a^2b+3a^3-ab^2+b^3$ take $5a^3+2ab^2-2b^3+4a^2b$ | | | | | 1. From the sum of $2a^3-a^2b+b^3$ and $3a^3+2ab^2-5b^3$ take the sum of $a^3+3a^2b-2b^3$ and $3a^3-ab^2+4b^3$ | | | | Subtraction |
| 2. Subtract $3x^3+2x^2y-y^3$ from $5x^3-2xy^2+y^3$ | | | | | 2. By how much does $3x-y+z$ exceed $2x+y-m$? | | | | |
| 3. By how much does $2x-y$ exceed $x+y$? | | | | | 3. By how much does 0 exceed $a+b$? | | | | |
| 1. Find the integer which is such that when increased by the first consecutive integer and this sum decreased by the second consecutive integer the result is 117. | | | | | 1. The sum of two consecutive odd integers is 72. Find them. | | | | Equations |
| 2. A sum of money amounting to \$3.90 consists of nickels and dimes. There are 41 coins in all. How many of each kind? | | | | | 2. One-fourth of a certain number exceeds one-sixth of it by 5. Find the number. | | | | |
| 3. $3(2x-4)(x+7)-2(3x-2)(x+5)=-3(x-2)$. Solve and check. | | | | | 3. A is 62 years old and B is 6 years old. In how many years will A be five times as old as B. | | | | |
| | | | | | 4. Solve $(x+10)(x-3)=(x-2)(x+6)$. Prove. | | | | |
| Divide: | | | | | Divide: | | | | Division |
| 1. $k^4+k^2m^2+m^4$ by k^2+km+m^2 | | | | | 1. $15x^3y^2-16x^4y+5x^3+13xy^4-11x^2y^3-6y^4$ by $x^3-3xy+2y^2$ | | | | |
| 2. $5x^2+\frac{1}{2}x-\frac{1}{2}$ by $3x-1$ | | | | | | | | | |
| Find prime factors of: | | | | | Find prime factors of: | | | | Factoring |
| 1. $9x^4-54x^2+81$ | 5. $ab-bx-ay+xy$ | | | | 1. $36a^2+36ax-40x^2$ | 5. $8ac-12bc-6af+9bf$ | | | |
| 2. $m^4a^4-81k^4$ | 6. $4x^2+8x-96$ | | | | 2. $36a^2+21ab-30b^2$ | 6. $36m^2-48a^2m+16a^4$ | | | |
| 3. $a^3b^3-a^4b^2+a^2b^4$ | 7. $24a^2+6ab-45b^3$ | | | | 3. $125k^3+64r^3b^3$ | 7. $81m^4n^4-16r^4$ | | | |
| 4. $64x^3-\frac{1}{y^3}$ | 8. $64m^3+r^3t^3$ | | | | 4. $15m^4a^3-6m^2n^4+3m^3n^3$ | 8. $8a^3-\frac{1}{27x^3}$ | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |
| | | | | | | | | | |

FIGURE 2—Continued

SECOND SEMESTER—ALGEBRA II

| TOPIC | C | B |
|------------------------|---|---|
| Factoring | Find prime factors of: 1. x^2+5x+6 2. x^2+6x+9 3. $x^2+xy+mx+my$ 4. $3x^2+5x+2$ 5. k^2-25n^2 6. $8-a^3$ 7. a^3+64 8. c^4-1 | Find prime factors of: 1. $ax-ay-kx+ky$ 2. y^4-10y^2+9 3. $9y^4-6y^2+y^2$ 4. $m^2-a^2+2ab-b^2$ 5. $27y^2+a^3$ 6. $c^3g^3-64c^3$ 7. $36y^2-36y^2$ 8. $36x^2-6x-12$ |
| Addition of Fractions | (Results must be expressed in lowest terms) Simplify: 1. $\frac{x-2a}{2ax} + \frac{2a-2z}{4az} + \frac{z-3x}{3zx}$ 2. $\frac{1}{x+4} + \frac{x}{1-x} + \frac{2x^2}{x+2}$ 3. $\frac{(x+7)(x-2)}{(x+4)(x+7)}$ 4. $m-a - \frac{m^2-a^2}{m+a}$ | (All results must be expressed in lowest terms) Simplify: 1. $\frac{m}{4(m+y)} + \frac{y}{3(m-y)}$ 2. $\frac{2a-b}{2a+b} + \frac{2a+b}{2a-b} + \frac{8ab}{4a^2-b^2}$ 3. $m^2+m+1 - \frac{1}{m-1}$ 4. $\frac{1-a}{1+a} - \frac{1-a^3}{1+a^3}$ |
| Fractional Equations | Solve and check: 1. $\frac{2}{3m} + \frac{3}{5m} = \frac{1}{12}$ 2. $\frac{m-5}{x+5} + \frac{m+4}{x+3} = \frac{69}{100}$ 3. $\frac{m-5}{4} + \frac{m+4}{5} = \frac{69}{100}$ | Solve and check: 1. $\frac{5x}{2x+5} + \frac{3x}{2x-5} = \frac{4x^2-80}{4x^2-25}$ 2. $\frac{x-2}{x-3} = 1 - \frac{3}{x-4}$ |
| Simultaneous Equations | 1. Plot $\begin{cases} 5x+y=7 \\ x-2y=-3 \end{cases}$ 2. Solve by Addition or Subtraction. $\begin{cases} 2x-10y=14 \\ 3x-5y=1 \end{cases}$ 3. Solve by Substitution. $\begin{cases} 3x+y=30 \\ 3y+x=34 \end{cases}$ | Solve: 1. $\begin{cases} \frac{x}{7} + \frac{y}{5} = \frac{10}{7} \\ x + \frac{y}{3} = \frac{14}{3} \end{cases}$ 2. $\begin{cases} \frac{9}{x} + \frac{4}{y} = 2 \\ \frac{18}{x} + \frac{18}{y} = 10 \end{cases}$ |
| Quadratic Equations | 1. Solve by completing the square $x^2-2x=4$ 2. Solve by factoring method $n^2-7n=18$ 3. Solve $x^2-49=0$ | (Results must be true to three decimal places) 1. Solve by completing the square $25x^2-20x=3$ 2. Solve by factoring method $2x^2-3x-5=0$ 3. Solve $x^2=5$ |

FIGURE 2—Concluded

SECOND SEMESTER—ALGEBRA II

| A | A+ | TOPIC |
|--|--|------------------------|
| <p>Find prime factors of:</p> <ol style="list-style-type: none"> $6x^2 - xy - 2y^2$ $6x^2 - 3ax + 2mx - am$ $1 - 64m^4$ $64k^3 + 8m^3$ $25a^4y^2 - 20a^2cy + 4c^2$ $16k^4 - 1$ $9x^2 + 3ax - 2a^2$ $x^2 - y^4x^2$ | <p>Find prime factors of:</p> <ol style="list-style-type: none"> $25m^3 - 9y^3 + 12ay - 4a^3$ $2x^3 - 8x$ $25m^4 + 50m^2a + 25a^3$ $6b^2c^3 - 25bc + 14$ $mx - mk + 2my - ax - 2ay + ak$ $y^4 + 3y^2 + 4$ $11x^2 + x - 10$ $27y^4 + 64m^3$ | Factoring |
| <p>(All results must be expressed in lowest terms) Simplify:</p> <ol style="list-style-type: none"> $\frac{4x}{(5+x)(4-x)} + \frac{5x}{(x+2)(x-4)}$ $\frac{b}{a-b} - \frac{a^2}{a^2+ab+b^2} + \frac{a^2-b^2}{a^2-b^2}$ $m^3+m+1 - \frac{m^3-2m+1}{m^2-m+1}$ $\frac{2m}{(a-5m)^2} - \frac{a-2m}{(a-3m)(a-5m)} + \frac{4}{a-3m}$ | <p>(All results must be expressed in lowest terms) Simplify:</p> <ol style="list-style-type: none"> $\frac{2+a^2}{a^3-9} - \frac{3a+1}{(3-a)(2-a)} - \frac{5a-1}{(a-2)(a+3)}$ $\frac{a+b}{a^2+ab+b^2} + \frac{a-b}{a^2-ab+b^2} + \frac{2b^2}{a^4+a^2b^2+b^4}$ $\frac{m^2-3m+9}{m^3-27} - \frac{m^2+3m+9}{m^3+27}$ | Addition of Fractions |
| <p>Solve and check:</p> $\frac{2x+1}{3x-12} - \frac{2x-1}{3x+15} = \frac{5x+3}{x^2+x-20}$ | <p>1. A has \$52 and B has \$68. After giving B a certain sum, A has left only one-half as much money as B then has. What sum did A give to B?</p> <p>2. A's age is one-third of B's age, and 6 years ago it was one-fifth of B's age. Find their present ages.</p> | Fractional Equations |
| <p>Solve:</p> $\begin{cases} \frac{6}{x+3} - \frac{3y+2}{3y+2} = 0 \\ \frac{1}{x+2} - \frac{2}{2y+2} = 0 \end{cases}$ <p>2. A is four-fifths of B's age; but in four years, A's age will be five-sixths of B's age. Find their present ages.</p> | <p>Solve for x and y:</p> $\begin{cases} bx - ay = b^3 \\ ax - by = a^3 \end{cases}$ <p>2. The sum of two digits of a number is 14; and if 18 be subtracted from the number, the remainder equals the number obtained by reversing the digits. Find the number.</p> | Simultaneous Equations |
| <p>1. Find two consecutive integers whose product is 240.</p> <p>2. The sum of a number and its reciprocal is $\frac{65}{25}$. Find the number</p> | <p>1. Find the base and the altitude of a triangle whose area is 432 square inches, if the base exceeds twice the altitude by 12 inches.</p> <p>2. Find the dimensions of a rectangle whose area is 720 square feet if the sum of the base and altitude is 54 feet.</p> | Quadratic Equations |

pupils who were barely passing had learned a great deal more than they really had, if measured in terms of absolute and not relative knowledge. Requirements had to be changed and tests rewritten to conform to classroom conditions.

After many trials, reducing the difficulty of one test increasing that of another, a set of tests was perfected as far as the limitations of two classes and two teachers would allow. The number of tests for each topic was fourteen, five C, four B, three A and two A+ tests. This number was found to be enough to cover all but the most exceptional cases, that is, no pupil who attempted five C tests and failed, was likely to succeed on the sixth.

The next endeavor was to show the value of the "Hurdle Scheme" in comparison with the traditional type of work. This was done in experimental and control groups. The test method easily proved its merit by the reduced number of failures and the increased knowledge of algebra. Although the C tests seemed ridiculously easy, the control groups could not pass them, showing that the knowledge of the pupils was not uniform or absolute but rather haphazard or hazy, that is they could do much more difficult work in which some principles were involved but could not do work which involved all of the principles demanded in the "minimum essentials" course.

The value in the use of "Hurdles" is also shown by a comparison of the final grades for groups in which the hurdles were used with those for the control groups in which they were not. Table I shows the distribution of grades D, C, B, A, and A+ among 63 pupils remaining to the end of the term in three sections in which the hurdles were used. Table II shows the same facts for 229 pupils in twelve sections in which the hurdles were not used.

Table III shows the percent of pupils receiving each grade in regular classes and in classes using the hurdles. The percent of failures is much greater among the regular classes as is also the percent of those receiving the lowest passing mark. A

TABLE I. PUPILS IN 3 CLASSES IN ALGEBRA II IN WHICH
 "HURDLES" WERE USED, CLASSIFIED BY GRADES RECEIVED—
 SCHOOL REPORTS, FIRST SEMESTER, 1915

| Class | Total | Number Receiving | | | | |
|----------|-------|------------------|----|----|----|----|
| | | D | C | B | A | A+ |
| I..... | 22 | 1 | 1 | 12 | 7 | 1 |
| II..... | 22 | 0 | 7 | 8 | 5 | 2 |
| III..... | 19 | 3 | 6 | 7 | 2 | 1 |
| Total | 63 | 4 | 14 | 27 | 14 | 4 |

TABLE II. PUPILS IN 12 CLASSES IN ALGEBRA II IN WHICH
 "HURDLES" WERE NOT USED, CLASSIFIED BY GRADES RE-
 CEIVED—SCHOOL REPORTS, FIRST SEMESTER, 1915

| Class | Total | Number Receiving | | | | |
|------------|-------|------------------|-----|----|-------|-------|
| | | D | C | B | A | A+ |
| I..... | 16 | 1 | 13 | 2 | | |
| II..... | 16 | 2 | 8 | 6 | | |
| III..... | 20 | 3 | 11 | 6 | | |
| IV..... | 18 | 1 | 7 | 10 | | |
| V..... | 23 | 1 | 14 | 6 | 2 | |
| VI..... | 18 | 1 | 12 | 5 | | |
| VII..... | 22 | 4 | 11 | 4 | 3 | |
| VIII..... | 20 | 3 | 8 | 5 | 2 | 2 |
| IX..... | 17 | 4 | 3 | 7 | 2 | 1 |
| X..... | 20 | 4 | 10 | 6 | | |
| XI..... | 19 | 4 | 11 | 4 | | |
| XII..... | 20 | 6 | 6 | 7 | 1 | |
| Total..... | 229 | 34 | 114 | 68 | 10 | 3 |

decidedly greater proportion of the pupils in classes using the hurdles received the higher ratings.

TABLE III. PERCENT OF PUPILS RECEIVING EACH GRADE IN REGULAR CLASSES AND IN CLASSES USING "HURDLES"

| Grade | Regular Classes | Classes using Hurdles |
|------------|--------------------|-----------------------------|
| 1 | 2 | 3 |
| D..... | 14.85 | 6.35 |
| C..... | 49.78 | 22.22 |
| B..... | 29.69 | 42.86 |
| A..... | 4.37 | 22.22 |
| A+..... | 1.31 | 6.35 |
| Total..... | 100.00 | 100.00 |

The facts of Tables II and III have been graphically combined in Figure 3. These graphs show the percent distribution by grades of the three classes using the hurdles and the twelve not using them. The near approach of Curve II to the normal and the skewing of Curve I to the right shows the advantage of the use of hurdles.

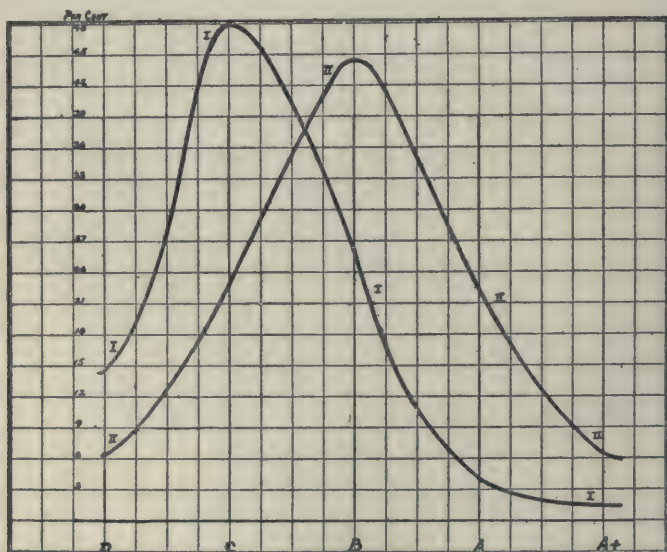


FIGURE 3. THE DISTRIBUTION OF GRADES IN THREE CLASSES IN WHICH "HURDLES" WERE USED, (CURVE II) AND IN TWELVE CLASSES IN WHICH THE "HURDLES" WERE NOT USED, (CURVE I).

TABLE IV. THE PUPILS IN ELEVEN CLASSES, ALGEBRA II, IN WHICH "HURDLES" WERE USED, CLASSIFIED BY GRADES RECEIVED—SCHOOL RECORDS—SECOND SEMESTER, 1916-1917

| Class | Total | Number Receiving | | | | |
|------------|-------|------------------|-------|----|----|-------|
| | | D | C | B | A | A+ |
| I..... | 19 | 3 | 3 | 6 | 6 | 1 |
| II..... | 20 | 2 | 7 | 7 | 4 | |
| III..... | 25 | 4 | 4 | 16 | 1 | |
| IV..... | 14 | 2 | | 9 | 1 | 2 |
| V..... | 17 | 2 | 3 | 3 | 7 | 2 |
| VI..... | 20 | 2 | 7 | 4 | 4 | 3 |
| VII..... | 21 | | 3 | 11 | 6 | 1 |
| VIII..... | 19 | | 6 | 4 | 8 | 1 |
| IX..... | 15 | 4 | 1 | 5 | 4 | 1 |
| X..... | 21 | 4 | 4 | 9 | 4 | |
| XI..... | 15 | 1 | 2 | 7 | 4 | 1 |
| Total..... | 206 | 24 | 40 | 81 | 49 | 12 |

Table IV shows the distribution of grades in eleven classes in which the hurdles were used after their introduction throughout the department. These classes and the twelve used in Table II were chosen because they were taught by the same teachers. While the results are not as satisfactory as those shown in Table I they are much better than those in Table II.

The percentage distribution is as follows:

| | |
|----------|-------|
| D | 11.65 |
| C | 19.42 |
| B | 39.32 |
| A | 23.79 |
| A+ | 5.83 |

| | |
|-------------|--------|
| Total | 100.01 |
|-------------|--------|

Figure 4 shows in graphic form the facts of Table IV. The divergance from normal lies in the fact that the curve is bent downward at the C and upward at the D line. Some of the pupils

who should have had C were evidently not reached and became D's.

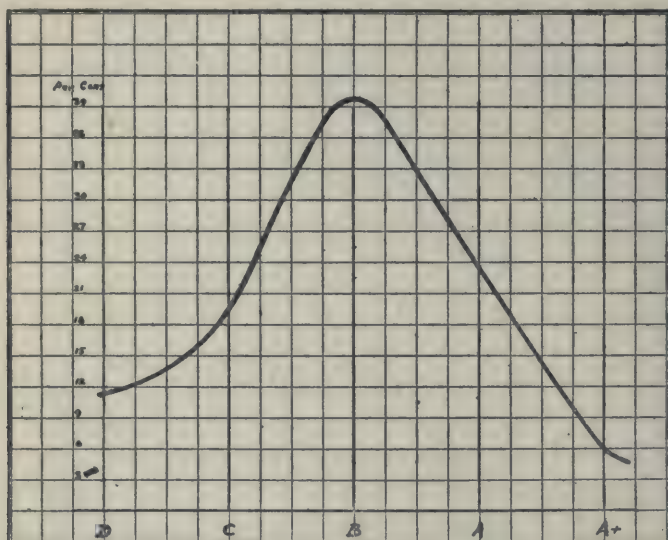


FIGURE 4. THE PERCENT DISTRIBUTION BY GRADES OF THE PUPILS IN ELEVEN ALGEBRA CLASSES IN WHICH THE "HURDLES" WERE USED

Table V shows for 16 classes of one semester the number gaining in rank or losing in rank in the successive topics. Under ideal conditions, with an ideal set of tests, the pupil who could achieve A+ in the first topic would achieve A+ in all topics and the pupil who was barely able to make C in one topic should be barely able to make C in all the others.

The tests are far from ideal and the conditions were only normal but the deviations in rank as pupils passed the successive tests were not many as evinced by the table. The tests on addition of fractions were slightly more difficult than any of the others because seven more lost than gained in rank and the tests on fractional equations were slightly less difficult because nine more gained than lost. Of the totals eight more gained than lost but this number is so small that as in the case of the gains or losses in particular topics, the irregularities of attendance, and interruptions would always cause that amount of deviation.

The discussion of methods has no place, theoretically, in any course of study, but, practically, the administration of a course

or the method of conducting the work is an integral part of the course. The "Hurdle System" was devised so as to interfere as little as possible with the usual methods of the recitation. It does, however, emphasize attention to individuals. On this account, therefore, a brief explanation of a change from the ordinary method may be desirable.

TABLE V. NUMBER OF PUPILS WHO GAINED OR LOST IN RANK IN CLASSES OF ALGEBRA II

| | |
|---|-----|
| Number of pupils in 16 classes remaining to end of term.... | 380 |
| Number making same grade in all tests of series..... | 240 |
| Number making gain in rank or loosing in rank or both.... | 140 |
| Number making gain in rank..... | 80 |
| In factoring..... | 18 |
| In addition to fractions..... | 18 |
| In fractional equations..... | 26 |
| In simultaneous equations..... | 18 |
| In quadratic equations..... | 7 |
| Number loosing rank..... | 72 |
| In factoring..... | 15 |
| In addition of fractions..... | 25 |
| In fractional equations..... | 17 |
| In simultaneous equations..... | 8 |
| In quadratic equations..... | 7 |

The teacher takes up any new topic as "Addition of Fractions" in any way he thinks best. As soon, however, as he thinks that one third or more of the class have a sufficient knowledge of principles and skill in manipulation, he gives a C test. He then has two groups and the difficulties of each child in the lower group are shown by his errors on the test. These points can be emphasized in any exercises the class is engaged upon, and class work goes on as usual. After an interval another C test and a B test are offered. Then there are three groups but the lowest group is very small and it is possible to give them individual attention and supplementary calibrated lists of exercises laying special stress upon the shortcomings of each pupil. The regular work, however, is continued, for the slow pupils can understand the solution of difficult exercises although they may not be skillful enough to manipulate the quantities successfully. On the one hand, the bright pupils are intent upon getting sufficient knowledge and skill to enable them to pass the higher tests which they

know will be offered as soon as they are ready for them. On the other hand, the slow pupils while they are doing tasks adapted to their skill are, at the same time, profiting by the teaching of the more capable pupils. For the latter are working on similar, though more difficult, material. In other words their exercises involve the same principles which the slower children are also attempting to apply.

Teachers soon develop a technic of handling pupils in groups, and since these groups are mechanically sorted out for their attention, they can arrange materials and methods to fit the needs of each group. This is especially easy since each group has advance notice or a sample of the requirements for the next test.

The chart on the wall keeps up the enthusiasm of every child and cures many cases of laziness because of the spirit of competition it engenders. The lazy pupil cannot hold the esteem of his fellows by an occasional show of brilliance and at the same time depend upon it to get him a good mark from the teacher at the end of the term. His record shows. He gets what he earns.

Reviews and Abstracts

TIDYMAN, WILLARD F. *The teaching of spelling*. (School monograph series.) Yonkers-on-Hudson, N. Y.: World Book Company, 1919. 178 pp.

The literature giving accounts of educational research which is appearing in increasing volume in our educational periodicals and in miscellaneous publications is making the demand for well-organized summaries a pressing one. An attempt to meet this matter has been made in the case of spelling by Dr. W. F. Tidyman in his recent monograph on the teaching of spelling. He states that his purpose is "to bring together from all sources the more reliable and pertinent facts of the teaching of spelling and to present them in their relation to the practical problem which the teacher has to face every day in the classroom." Aside from the fact that teachers and superintendents generally do not have access to the original sources, it is very difficult task to translate the results of research into guiding principles and rules for schoolroom practice. Therefore, the writer of this monograph has rendered distinctive service in bringing together some of the more reliable and pertinent facts in the field of spelling.

The table of contents is a good index of the topics treated:

CHAPTER

- I. The Selection and Classification of Words
- II. Preliminary Testing for Word Difficulty
- III. The Psychological Basis of Spelling
- IV. The Presentation of Words
- V. Independent Study and Reviews
- VI. The Prevention and Treatment of Errors
- VII. Testing
- VIII. The Measurement of Spelling Efficiency
- IX. Factors Affecting Spelling Efficiency

APPENDIX

- A. Spelling Plans
- B. Minimum Word List

BIBLIOGRAPHY

Spelling Material
The Psychology of Spelling
The Pedagogy of Spelling
Scales and Standards Tests
Factors Affecting Spelling Efficiency
General Studies—Summaries, Discussions, etc.

One of the most valuable features of the book is the bibliography. In addition to being classified under the headings given above, there are help-

ful subdivisions. For example, under the general heading of spelling material, the sources are grouped under the following sub-topics:

Word lists, difficult words, common misspellings, classes of errors, grading, grouping, homonyms.

Under the pedagogy of spelling we find the following subdivisions:

Syllabification and diacritical marks, meaning, divisions, method of presentation, spelling plans, the value of rules, transfer, class study versus independent study, spelling consciousness, correction or errors.

Such a grouping of references for particular topics will be very helpful to anyone who wishes to consult the original sources.

A net total of 114 sources is mentioned. Only a very limited number of unpublished theses are included. It is probably too much to expect of a writer to include in such a summary such sources but there is a considerable accumulation of unpublished reports some of which are very valuable. For the most part, they have not been listed and extensive and persistent correspondence would be required to unearth them. In comparison with some other bibliographies on spelling, it appears that sources in the general field of psychological research have not been completely canvassed. This shortcoming is, however, offset by the inclusion of a number of references from foreign publications.

The style of the book is simple and teachers will find it readable. Few direct citations to sources are made and interwoven with the summary of the results of research is much of the writer's pedagogical theories and beliefs in this field. Use is made of paragraph side headings in italics as well as frequent centered headings. This makes it easy to locate in the chapter the discussion of even minor topics. At the end of each chapter there is a summary consisting of numbered statements which are in the form of principles and rules for the teaching of spelling.

Students in the field of education will be disappointed that the writer has not given specific citations to the sources on which his principles and rules are based. However, this omission will not be missed by the general reader. In general, the writer does not discuss in a critical way the accounts of research which he uses. However, a critical comparison of analyses would not be in keeping with his announced purpose. The book will meet a real need and Dr. Tidyman is to be commended upon his effort.

W. S. M.

LYON, LEVERETT S. *A survey of commercial education in the public high schools of the United States.* (Supplementary Educational Monographs, V. 3, no. 5, 1919.) Chicago: University of Chicago Press. 62 pp.

Commercial education in high schools has been slow to accept scientific procedure in its reorganization. Developing as it did immediately out of the practices of business colleges, commercial education has been more or less formal and routine in character. One of the first steps in a progressive modification of traditional practices is a thorough-going study of existing practices, together with a careful analysis and evaluation of these practices in the light of modern social needs.

Mr. Lyon's monograph reports the results of a careful investigation of the practices in the commercial departments of one hundred thirty-six high schools in different sections of the country. Among the problems which were studied are the following: the percentage of students in commercial courses, the length of commercial courses, the kinds of subjects included in these courses, and the English requirements. Special consideration is given to the "social-business" subjects. The fact is emphasized that there is genuine need for agencies through which much-needed text and reference material can be organized and published. In a final chapter entitled "Plans and Policies in Secondary Commercial Work," Mr. Lyon points out the need for a study of the purposes of high school commercial courses in "modern functionalized society." In addition, he shows the need of scientific surveys and investigations to secure the facts on which to base radical changes in current practices.

Although Mr. Lyon's study is by no means exhaustive, it represents the most comprehensive and scientific investigation of its kind which has been made in the field of commercial education. It should be read by superintendents, high school principals, and all teachers of commercial education.

W. S. G.

HOBSON, ELSIE GARLAND. *Educational legislation and administration in the state of New York from 1777 to 1850*. (Supplementary educational monographs, v. 3, no. 1, 1918.) Chicago: University of Chicago Press. 267 pp.

In this monograph, Miss Hobson has analyzed and summarized the educational legislation of the state of New York from 1777 to 1850 in terms that are intelligible to the average reader. The following topics receive considerable emphasis in her discussions: formative influences; dual system of school control; education under regents; the common school system; special legislation for cities; support of education; and education of special classes. The appendix contains complete lists of titles of the legislative acts with exact references to the sources where each may be found; abstracts of the charters of colleges, academies, and other types of institutions; abstracts of legislation concerning school lands and funds; common school acts and other similar data.

Miss Hobson has investigated a field in which much research work is necessary. It is obvious that an exact knowledge of educational legislation during the early periods of our history is necessary, if we are to understand the evolution of the various public school systems. It is also evident that this type of educational history will supply facts for interpreting and modifying present school practice. Furthermore, it will aid greatly in throwing light on our social, religious, economic and political history. From various points of view, therefore, this monograph is a valuable contribution to the history of American education.

W. S. G.

PUBLICATIONS RECEIVED

- Baylor, Adelaide Steele, and Colbert, Emma. *Young America's First Book*. New York: Century Company, 1919. 167 pp.
- Betelle, James O. *Checking Schedule for Projected School Buildings*. Milwaukee: Bruce Publishing Co., 1919. 32 pp.

- Betts, George Herbert. *How to Teach Religion, Principles and Methods*. New York: Abingdon Press, 1919. 223 pp.
- Bexell, J. A. *First Lessons in Business*. Philadelphia: J. B. Lippincott Co., 1919. 174 pp.
- Bronner, Augusta F. *The Psychology of Special Abilities and Disabilities*. Boston: Little, Brown and Company, 1919. 269 pp.
- Chamberlain, Arthur Henry and Chamberlain, James Franklin. *Thrift and Conservation, How To Teach It*. Philadelphia: J. B. Lippincott Co., 1919. 272 pp.
- Cubberley, Ellwood P. *Public Education in the United States*. Boston: Houghton Mifflin Co., 1919. 517 pp.
- Dewey, Evelyn. *New Schools for Old*. New York: E. P. Dutton and Co., 1919. 337 pp.
- Elson, J. C. and Trilling, Blanche M. *Social Games and Group Dances*. Philadelphia: J. B. Lippincott Co., 1919. 258 pp.
- Elson, William H. and Keck, Christine M. *Junior High School Literature*. Chicago: Scott, Foresman and Co., 1919. 624 pp.
- Field, Walter Taylor. *Readings from English and American Literature*. Boston: Ginn and Co., 1919. 612 pp.
- Guiteau, Wm. Backus. *Our United States*. New York: Silver, Burdett and Co., 1919. 680 pp.
- Hamilton, Samuel. *Essentials of Arithmetic, Book I*. New York: American Book Company, 1919. 392 pp.
- Hamilton, Samuel. *Essentials of Arithmetic, Book II*. New York: American Book Company, 1919. 460 pp.
- Haney, James Parton. *Industrial Art, Its Tenets and Teaching*. New York: Prang Company, 1919. 37 pp.
- Jackson, Bennett B., Deming, Norma H., and Bemis, Katherine I. *Thrift and Success*. New York: Century Company, 1919. 288 pp.
- Laing, Graham A. *An Introduction to Economics*. New York: Gregg Publishing Company, 1919. 454 pp.
- Lefferts, Walter. *American Leaders, Book I*. Philadelphia: J. B. Lippincott Co., 1919. 329 pp.
- Lefferts, Walter. *American Leaders, Book II*. Philadelphia: J. B. Lippincott Co., 1919. 354 pp.
- Lemos, Pedro J. and Lemos, John T. *Art Simplified*. Chicago: Prang Company. 174 pp.
- Lewis, Charles D. *The Waterboys and Their Cousins*. Philadelphia: J. B. Lippincott Co., 1918. 172 pp.
- McDonald, Etta B. and Dalrymple, Julia. *Umé San in Japan*. Boston: Little Brown and Co., 1919. 119 pp.
- McLaughlin, Andrew C. and VanTyne, Claude Halstead. *A History of the United States*. New York: D. Appleton and Co., 1919. 575 pp.
- McPherson, William L. *The Strategy of the Great War*. New York: G. P. Putnam's Sons, 1919. 417 pp.
- Mahoney, John J. *Standards in English*. Yonkers: World Book Co., 1919. 198 pp.
- Mirick, George A. *Home Life Around the World*. Boston: Houghton Mifflin Co., 1918. 158 pp.

Editorials

SCIENTIFIC SUPERVISION

Supervision is a term whose meanings range from the general oversight which a superintendent gives to the school system as a whole to the aid given by one teacher to another. However, supervision of instruction is a well defined and important phase of educational activity in American schools, and as such is a proper field for measurement and research. Moreover, it has come about in the chaotic and uncontrolled evolution of our school system, that in the past little attention has been directed to the nature and laws of supervision so that the opportunities for constructive efforts in this field are very great.

This Journal will welcome accounts of the application of scientific methods to supervisory problems. School men everywhere are beginning to consider the best ways of organizing supervision of instruction is a well-defined and important phase. They are analyzing the work of supervision and finding that it consists of many different types of activity. The inspection and rating of teachers is a well-recognized function of supervisory work and one which eventually entails measurement. Instruction and assistance are two other functions in which measurement may be of very great aid, both in selecting the teachers in need of training and in evaluating the effect of the training given. Then there are the larger problems of curriculum building, the setting of standards, the determination of the efficiency of various methods. Finally, there is the problem of determining the effect of supervision itself and of the condition under which it produces its optimum results.

Scientific methods are already rapidly transforming supervision and will greatly increase its effectiveness. Any school system doing constructive work in this field should feel it a duty to publish an account of its experiments for the stimulation and guidance of other workers. Contributions of this character are solicited. From time to time, also, articles dealing with the

general principles and theory of supervision will appear in these pages, and every effort will be made to make the Journal of service to those whose interests center in supervisory activities.

S. A. C.

EDUCATIONAL TESTS AS IMPORTANT SCHOOLROOM EQUIPMENT

During the decade which is just closing much emphasis has been given to demonstrating to schoolmen that the results of teaching can be measured by scientifically devised instruments. Within this period a large number of educational tests have been devised and several of them have been and are now being extensively used. It appears that there is no longer any need to demonstrate the possibility of measurement. It is accepted for certain school subjects by even the conservatives. The need today is rather for educating superintendents, principals, and teachers in the effective use of educational tests. This need is particularly urgent in the case of teachers for two reasons. First, there are many more teachers than superintendents and principals and in general they have not had the same opportunity of learning about educational tests. In this connection, it may also be mentioned that a very large percent of teachers enter the profession each year. Second, the most important reason is that educational tests are instruments which teachers may use to increase the effectiveness of their instructional efforts. But in order that the tests may fulfill this function, it is necessary that teachers know how to use them. To give them is merely to take the first step. It must be followed by interpretation of the scores in terms of pupil needs and modification of instruction in accordance with these needs.

This use of educational tests for the improvement of instruction is not incompatible with their use for supervisory purposes, but the writer believes this supervisory use should not be emphasized more than it is now being emphasized. In fact, there are cases on record where this function has been over-emphasized. This tends to prejudice teachers against the use of educational tests or at least to divert their attention from the instructional function.

This department of the JOURNAL OF EDUCATIONAL RESEARCH will emphasize the use of educational tests for the improvement of instruction. It will welcome communications from teachers who have found the test helpful. Those who have secured increased results from their teaching owe it to the profession to make known their experience. Even if it appears commonplace, it may contain just the suggestion which other teachers need.

The use of educational tests as instruments for the improvement of instruction places them in the list of important school-room equipment. We consider it necessary for a teacher to be supplied with blackboards, crayon, maps, reference books, supplementary readers, and other similar equipment in order that he may render efficient service. On exactly the same basis, it is necessary that the teacher be supplied with appropriate educational tests if his instructional efforts are to be most effective. A good teacher will, of course, secure commendable results without them, but so will a good teacher secure commendable results without supplementary readers, reference books, and the like. Many boards of education are now purchasing educational tests for use throughout their school systems. In fact, it is the exception rather than the rule when the teacher is required to pay for the tests. But definite provision for the purchase of testing materials should be made in the school budget. Until this is done, we cannot say that they have official recognition as important schoolroom equipment.

W. S. M.

ON THE ORGANIZATION OF BUREAUS OF RESEARCH

The membership roll of the National Association of Directors of Educational Research at this writing carries sixty-seven names (includes five honorary members) and, geographically, stretches from the Atlantic to the Pacific. This Journal as an official organ will publish from time to time a full account of the business and activities of the Association as well as special reports and news items dealing with the work of particular bureaus in various cities.

It is probable, however, that the Journal might render valuable service of a little different type. Research work in many school

systems has not yet grown to a magnitude which warrants the establishment of a separate department, while in many other systems newly established bureaus are struggling and staggering through the trying "first years" of independent existence. Problems and difficulties without number are sure to arise, and too often the person responsible for "putting measurement across" in his school system must meet these difficulties and misunderstandings alone. Perhaps he is the only person in the school system who really understands what educational research means, who really has a scientific attitude of mind towards educational problems. What such a person needs in the first place is the direct, personal sympathy and support of others engaged in the same work and struggling with the same difficulties. Then, in the second place, he needs direct and immediate help from those who have come through the fight successfully and who know how the thing is done.

It is the wish of the associate editor in charge of the work of bureaus of research that his department may minister to both these needs. He will enjoy an exchange of personal, confidential notes from fellow-workers and will strive to bring those whose difficulties are similar together. He will do what he can to answer questions and give advice, or to pass the questions on to those who can supply the answers. Particularly he will welcome suggestions concerning helpful articles for publication on the organization and activities of bureaus of research. With a field so large to choose from he is keenly conscious of the difficulties of making a right selection and appeals to all true friends of educational measurement to guide him in his editorial duties.

S. A. C.

National Association of Directors of Educational Research

(E. J. ASHBAUGH, *Secretary*)

This department will be devoted to the affairs of the National Association of Directors of Educational Research. Here news and comment concerning the members of the Association and their work will appear. Official announcements, reports of meetings, statements of committees, names of new members, and many less formal matters will find an appropriate place. It is planned that the "News Letter," which has hitherto been privately circulated among the members of the Association, shall be given wider publicity in these pages. As secretary of the Association, Professor E. J. Ashbaugh, Extension Division, University of Iowa, Iowa City, Iowa, is a member of the editorial board of the Journal and will edit this department. Communications having reference to the department should be addressed to him.

B. R. B.

ANOTHER STEP FORWARD

This is a time of advancement, of progress, of facing forward. The close of the great war which had speeded up production in many lines and compelled men to think in new terms, to make new adjustments, and, in a sense, move at a more rapid rate, did not result in everything settling back into the pre-war condition. While there has been decreased activity in some lines, perhaps even stagnation, in the realm of educational thought, the reaction has not set in—at least not yet.

At the Cincinnati meeting of the Department of Superintendence, the National Association of Directors of Educational Research was founded. On that occasion eight men met at a luncheon by an almost impromptu arrangement. These men were Ayres, Ballou, Buckingham, Courtis, Hebden, Melcher, O'Hern, and Shiels. This was in February 1915, less than five years ago. At the Detroit meeting the following year, the first program was put on and the first banquet was held.

A year later a larger meeting and a larger group at the banquet table evidenced the fact that the newly organized association was attracting the interest of the men who were actively engaged in this field of work. Not only so, but the leaders were ambitious and agreed at that time to produce a volume for the National Society for the Study of Education. One result of that decision was Volume II of the Seventeenth Year Book. How far-reaching that decision has been no one can tell. That the Year Book has had a tremendous influence in directing

the thought of superintendents, supervisors and teachers toward educational measurement, cannot be doubted.

At the same meeting it was decided to circulate a bulletin among the members as a means of keeping in touch with the activities of each other. This bulletin is now rounding up its third year. That it has been helpful to the various members was evidenced a year ago when a membership fee was first levied in order that the bulletin might be continued.

Another step has now been taken by the Association. By an almost unanimous vote the organization decided to adopt the new JOURNAL OF EDUCATIONAL RESEARCH as its official organ. This journal is now making its bow to the public, and the Association is presenting itself to a larger field than ever before through this department of the Journal. To what extent the Journal will be beneficial to the organization remains to be seen.

It is highly important that with this new venture, each member of the Association shall resolve to continue to advance. The high ideals of the founders of the Association have been partly realized. The dreams of those most deeply interested in it are continually expanding. That our work is important, we know; that the results have been, and are, valuable in our own local situations, has been demonstrated; that the individual members of the Association may make their work beneficial to an ever enlarging group, is at least one of the purposes in the adoption of the Journal.

A large number of our members have been made contributing editors. The Association, through its secretary, is represented on the editorial board. Short news items concerning the things the members have in progress, and the problems they have in mind, should be sent to the secretary frequently. In fact, the best means by which each individual member may contribute to the success of the Association as a whole, is by notifying the secretary the moment a new idea or a new plan is conceived.

ACTIVITIES OF THE MEMBERS

Practically all directors of city systems report the giving of one or more pedagogical tests. An examination of the tests used quite uniformly shows Courtis, Series B, and some form of reading, writing and spelling.

The following members report the use of mental tests during the current year:

Brainerd, Mrs. Margaret, Director of Department of Reference and Research, Martin's Ferry, Ohio.

Dalman, Murray A., Director of Department of Reference and Research, Board of Education, Indianapolis, Indiana.

Dickson, Virgil E., Acting Director of Reference and Research, Public Schools, Oakland, California.

Henry, Mary Bess, Director of Research, Santa Ana, California.

Lane, Robert E., Division of Educational Research, Los Angeles, California.

Lefler, M. C., Director of Measurements and Research, Lincoln, Nebraska.

Race, Henrietta V., Director Psychological Clinic, Board of Education, Louisville, Kentucky.

Smith, Leon O., Director Educational Research, Omaha, Nebraska.

Surveys of local schools, and more or less state-wide testing on certain subjects are reported from the following state bureaus:

University of Iowa, Iowa City, Iowa.

University of Illinois, Urbana, Illinois.

State Normal School, Emporia, Kansas.

University of Nebraska, Lincoln, Neb.

State Department, Madison, Wis.

University of Washington, Seattle, Wash.

Progress in the work of standardization of new tests is reported by:

Dr. B. R. Buckingham and Dr. Walter S. Monroe, University of Illinois.

J. C. De Voss, State Normal School, Emporia, Kansas.

Dr. W. W. Theisen and Cecile White Flemming, State Department, Madison, Wisconsin.

Two members—Leon O. Smith of Omaha and Virgil E. Dickson of Oakland—report the interesting experience of each passing through a five million dollar school bonding issue. Reports on the result have not yet been received.

Many interesting problems are being attacked by various members. Among those listed in recent letters are the following:

A study of the ability of elementary, high school and college students to discriminate between various qualities of poetic expression.

The educational value of motion pictures.

The influence of the make-up of a book upon the child's desire to read it.

Oral presentation vs. written presentation vs. graphic presentation.

The value of supervision with the use of tests vs. supervision without the use of tests.

The study of a few industries for the sake of learning the requirements of those industries into which the majority of our school children go as workers.

Devising tests of the fundamentals in music as taught in the primary grades.

Does departmental teaching result in better teaching of mathematics?

To what extent do teachers whose class evercises are aparently of superior merit secure better results as measured by subject matter tests?

The complete program of the Thursday meeting at Cleveland will be printed in the next issue of the JOURNAL OF EDUCATIONAL RESEARCH. Plans for the closed meeting on Monday will also be announced. The holding of a closed meeting is a departure from previous practice. Attendance will be limited to Association members and their invited guests.

The idea of having a headquarters hotel at Cleveland seems to have met with general approval. The following members have notified the president that they have reserved rooms at the Hotel Statler:

| | | |
|------------------|-------------------|----------------------|
| Murray A. Dalman | C. W. Flemming | Samuel B. Allison |
| James C. Devoss | Henrietta S. Race | Charles H. Elliott |
| C. E. Chadsey | Carter Alexander | W. W. Phelan |
| Leon O. Smith | Clara L. Mallory | Harriett M. Barthel- |
| William S. Gray | Ira J. Bright | mess |
| W. W. Theisen | W. E. Haggerty | Arthur W. Kallom |
| Walter S. Monroe | S. A. Courtis | E. L. Thorndike |
| F. E. Shapleigh | Chas. Fordyce | J. W. Richardson |
| E. J. Ashbaugh | F. J. Kelly | Charles H. Judd |
| Paul H. Hanus | Robt. S. Lane | H. A. Brown |

All new members and such old members as have not reserved rooms at the headquarters hotel should do so at once.

It appears that the addresses presented at our last meeting are not to be printed in the Proceedings of the N. E. A. The Proceedings are being pretty well filled up with addresses and papers from sections of the N. E. A. and from affiliated societies whose age is greater than ours. It is fortunate, therefore, that we now have a new means of publication. The JOURNAL OF EDUCATIONAL RESEARCH offers its pages for the purpose. The secretary gives notice at this early date that those who are to appear on the program at Cleveland will be requested to furnish him with a copy of their manuscripts.

News Items and Communications

Meetings of the State School Board Association and the City Superintendents' Association of Illinois were held at Urbana, Illinois, on October 29 and 31, 1919.

Superintendents and Board Members

These two associations have been meeting together now for four or five years. By thus combining their annual meetings they have in many ways brought about favorable results for the schools.

Working together the two associations had a large part in securing the passage of the Hicks' Law, providing larger revenues for the schools. One of the chief objects of these meetings is to create in board members a consciousness of the importance of the work they are undertaking. Another is to enable the superintendents to conduct their work more successfully through the better understanding that these joint meetings permit between the professional and lay officers of the schools.

The two associations met jointly during the afternoon and evening of the twenty-ninth and during the morning of the thirtieth. During the afternoon of the thirtieth the associations met separately.

At the first joint session the general subject was Physical Education. Assistant Superintendent Ambrose B. Wight of Chicago proposed a definite place for physical education in each day's program, specially prepared supervisors and teachers, adequate equipment, and adaptation of the work to pupils of different ages and maturities, regardless of their grading in the schools.

Mr. Charles Krauskopf, principal of the Victor-Lawson School of Chicago, described the origin of the physical education movement at Chicago, showing its growth from formal gymnastics into directed play. He told of the interesting experiment that is being tried in his own school in organizing play activities. He stressed the idea that any program adopted should be educational in a broad sense—that is, should make provision for all the pupils and develop not only physique but also character. Mr. Krauskopf believes that this type of physical education affords unrivaled opportunities for the development of desirable traits of character.

Superintendent R. O. Stoops of Joliet told about the experiment which is being tried out in that city. A co-ordinating board consisting of the superintendent of the elementary schools, the principal of the township high school, the mayor of the city, and a representative of the parochial schools has been formed. By the use of additional money voted at a recent election, the Park Board is to furnish a recreational

director for the city. The specific activities are to be financed by the high school, the elementary schools, the parochial schools, and city parks.

At the evening session on October 29, the subject was "Americanization." Professor Thomas A. Moran of Purdue University gave a fine inspirational address on this subject. He urged that the American people should actively take up their responsibilities to the immigrant.

Superintendent C. R. Reed of Rockford responded with a short talk in which he mentioned somewhat more specifically plans to be taken up among foreign people. The discussion was closed by Professor J. W. Garner of the University of Illinois, who in a brief but vigorous address urged the conservation of all that is good in our institutions, and the protection of these institutions from the alien agitator.

At the close of the evening program, Dr. B. R. Buckingham, Director of the Bureau of Educational Research, presented a plan for organizing a Superintendents' Conference under the auspices of the University. The project contemplates the meeting of such a conference annually during the same week as that in which the High School Conference meets. Subsequently at its business meeting the Superintendents' Association appointed a committee to consider what action the Association should take in relation to this proposal. The committee is to report at a meeting to be held at Springfield during the week between Christmas and New Years.

The topic of the meeting the following morning was "The Training of Teachers before and during Service." Superintendent John W. Withers of St. Louis, Missouri, gave the principal address. He warned of the danger of deterioration in our schools because of the lack of high class recruits for the teaching force. He urged the need of better salaries, of greater dignity for the teaching profession, and of more effective leadership.

He saw the necessity of inducing high class young men and women to enter our training schools. He summarized plans that had been worked out in St. Louis for extension courses for teachers already in service.

Dean C. E. Chadsey, of the College of Education of the University of Illinois, discussed the topic, urging the need of higher salaries and the value of securing college graduates as elementary teachers. The topic was further discussed by Dr. E. H. Taylor of the Charleston State Normal School who suggested among other things, the sending out of training teachers from the normal schools into the public schools to supervise teaching; by President J. Stanley Brown, of the DeKalb State Normal School, who pointed out that the normal schools of the state are now handling about 60 percent of their possible attendance and urged the desirability of getting more young women to enter these schools; and

by Superintendent T. E. Savage of Havana, who considered some of the plans specifically adapted to the smaller cities.

During the afternoon of October 30 the School Board Association was addressed by Mr. W. G. Spurgin of Urbana on the subject of "Consolidation of School Districts and the Enlargement of the Territorial Unit." After the address a business session was held.

At the same time the City Superintendents' Association, prior to the business meeting, listened to an address by Miss Ida Hoffman of the East St. Louis public schools on "The Teacher's Viewpoint of Constructive Supervision." She believed that all supervision should be constructive, that teachers should always be invited to comment on suggestions offered and to offer solutions for problems in supervision, and that supervisors should encourage teachers at each step in their work by pointing out "high spots." This paper was discussed by Superintendent H. A. Perrin of Jacksonville, Illinois, who urged the importance of standards of achievement, and by Superintendent W. R. Curtis of Kewanee, Illinois, who emphasized the participation of teachers in supervision.

By having spelling tests given in a number of states, Dr. B. R. Buckingham secured data which made it possible to add 505 new words to the 1,000 words of the Ayres Spelling Scale. With the very generous permission of Dr. Ayres, this list together with the original 1,000 words is now published under the title **Ayres Spelling Scale Extended** of Buckingham's Extension of the Ayres Spelling Scale. The new words tended to be more difficult than those of the original scale. This was felt to be fortunate because the Ayres Scale is less useful for upper grades than for middle and lower grades because of the fewness of the words at the high end of the scale.

It should be understood, however, that the extended scale does not in its entirety constitute a fundamental vocabulary scientifically determined. The added words were not obtained, as were the original words, from tabulations of frequency of use. They were selected from a much larger list secured by taking the words common to a number of spelling books. Therefore, one should not assume that the 505 new words are the next most common words after the 1,000 contained in the original Ayres Spelling Scale.

In constructing the Extension, the words of the original scale were not disturbed. All of the new words are placed in the appropriate columns under the old words and printed in italics. The new words have thus been kept separate from the old ones and a person who secures the Extension also has the original Ayres Scale intact.

Buckingham's Extension of the Ayres Spelling Scale is published by the Bureau of Educational Research of the University of Illinois. Sample copies will be sent postpaid on receipt of 10 cents. In quantity the Scale will be furnished for 8 cents per copy.

The first program meeting of the Illinois Section of the Mathematical Association of America was held at the University of Illinois on November 22, 1919. There were fifty-nine persons present which included some visiting high school teachers of mathematics. The program was as follows:

1. Address by H. E. Slaught, President of the Mathematical Association of America.
2. Round Table Discussion of the Topic: "Freshman Mathematics as Related to Varying Admission Credits from the High Schools." Discussion led by C. E. Comstock, Bradley Polytechnic Institute and M. W. Coultrap, Northwestern College.
3. Address: "The Training of Mathematics Teachers," by E. B. Lytle, University of Illinois.
Discussed by Malcolm McNeill, Lake Forest College, Bessie I. Miller, Rockford College, W. J. Risley, James Millikin University.
4. Address: "How Mathematicians Work," by Henry Blumberg, University of Illinois.

President Slaught outlined the history of mathematics teacher organizations in the United States and led up to the special need and great opportunities for real service facing this new organization.

The discussion of high school courses and college entrance credits was lively and suggestive of needed investigations and standardization. There seemed to be a general feeling that skill in formal symbolic manipulation was being overemphasized to the detriment of mathematical thinking in problem situations.

Mr. Lytle outlined a course of training desirable for all high school teachers of mathematics but remarked that such standards of preparation were hardly worth considering until increased salaries were sufficient to attract our ablest men and women into the teaching profession.

Mr. Blumberg gave some concrete ways in which some great mathematicians had worked and made a few generalizations in the way of suggestions for more effective methods of work.

The interest exhibited throughout this meeting promises well for the future of this new organization. It is the intention to hold annual meetings; and there is no reason why valuable contributions cannot be made to investigations in curriculum and teaching problems in both high school and college mathematics.

The former officers were elected for another year. They are:

President: J. A. Foberg, Craine Junior College.

Secretary: E. B. Lytle, University of Illinois.

Realizing in an indefinite way the undoubted value of many reports of investigations as these have been embodied in master's and doctor's

dissertations, the director of the Bureau of Educational Research of the University of Illinois, undertook to obtain from the Graduate Theses Listed heads of the departments of education in the various colleges and universities of the country a list of the authors and titles of such dissertations as have been accepted since January 1, 1917. The attempt was made more or less as an experiment. Unusual interest, however, was manifested by those who sent in material and it has been thought wise to assemble it as an informal report. It can hardly lay claim to completeness, although all the heads of departments addressed responded either by indicating that they had nothing to report or by furnishing the desired lists.

A classification of the dissertations will be made so far as the titles make this possible. It is estimated that the material will constitute about seventy mimeographed pages. A copy of the report will be sent to all those who contributed to it. A limited number of reports will also be offered for sale at one dollar each. These dissertations have sometimes been printed either completely or in the form of abstracts. These facts have been noted in the report as far as known. Even unpublished dissertations may generally be borrowed for an individual by a library.

The Extension Division of the University of North Dakota has a Bureau of Educational Cooperation which includes the activities of the State High School League. At the present time these are debate, declamation for boys and for girls, music in its various forms, and athletics. A state contest is held at the University in May each year. At the same time the county superintendents and high school principals have a three day conference. The director of extension work, Albert H. Yoder, attempts to answer professional questions arising in connection with school work. He has also organized a child welfare department. The medical school, state public health laboratories, departments of education and psychology, and the university library are cooperating in this work. A child welfare clinic was opened December first. Another new feature of the division is the University Extension Reading Course intended to supply a service to the people of the state desiring such intellectual opportunities as can be realized during the long winter evenings. The list this year contains about a dozen titles in as many fields, as:

Thayer, William Roscoe. *Theodore Roosevelt; An Intimate Biography*. Boston: Houghton Mifflin Co., 1919.

March, Francis Andrew. *History of the World War*. Philadelphia: J. C. Winston Co., 1918.

Gompers, Samuel. *American Labor and the War*. New York: Geo. H. Doran Co., 1919.

Ogg, Frederick Austin. *National Progress 1907-17 (American Nation a History, V. 27.)* New York: Harper and Bros., 1918.

Foster, John Watson. *Practice of Diplomacy as Illustrated in Foreign Relations of the United States*. Boston: Houghton Mifflin Co., 1906.

Trowbridge, Edward Dwight. *Mexico Today and Tomorrow*. New York: Macmillan Co., 1919.

The Executive Committee of the Wisconsin State Teachers' Association this year consented to add an Educational Measurements Section to its already long list of sectional meetings. Dr. W. W. Theisen, State **An Educational Research Section** Director of Educational Measurements, arranged the program. The meeting drew an attendance of more than five hundred even though scheduled at the same hour as several of the stronger sections. Following the program a number of the leaders met at a luncheon at the Hotel Wisconsin for the purpose of effecting a permanent organization of educational research workers. The new organization will be known as the Wisconsin Educational Research Association. Its purpose is not only to foster educational research but to keep members in touch with what other members are doing in the way of educational investigation and experimentation.

Officers elected were: Superintendent C. J. Anderson of Stoughton, President; Mrs. C. W. Flemming, State Supervisor of Educational Measurements, Secretary-Treasurer. Dr. Edgar F. Riley of the Platteville Normal was chosen president of sectional meeting for next year.

At the last session of the Legislature of North Carolina an appropriation was made for conducting a State Survey. Dr. Frank P. Bachman, of the General Education Board has been appointed to direct the work. Professor M. E. Haggerty of the University of **North Carolina Survey** Minnesota will have immediate charge of the tests and measurements. Examinations will be given in the high schools and in the elementary schools of both rural and urban districts. The present plan contemplates giving tests in spelling, arithmetic, reading, algebra, Latin, and English composition. It is proposed also to submit intelligence tests—at least in a limited way. Professor Haggerty will go to North Carolina about January first to take up his part of the work of the survey.

An appropriation of fifty thousand dollars in five annual payments has been granted to the Iowa Child Welfare Research Station of the State University of Iowa, of which Dr. B. T. Baldwin is director, by **Child Welfare Research Station** the National Women's Christian Temperance Union, for the purpose of making detailed scientific investigation into the development and conservation of child life. The gift is for research and the station will be given a wide range of freedom in the selection of its problems, which will center around the field of eugenics, nutrition, and the mental and social development of the child.

IMPROVEMENT IN SCHOOL SUBJECTS THROUGHOUT THE SCHOOL YEAR

V. A. C. HENMON
University of Wisconsin

The nature, amount, and rate of improvement in the case of relatively simple functions has been studied in some detail. The more complex abilities in school subjects have scarcely been touched. In discussing improvement in informational, appreciative, analytic, and selective functions, Thorndike pertinently remarks: "It is a noteworthy commentary on the newness of scientific study of education that nobody has yet measured the actual progress of any single child in learning any school subject for over a month, or in learning any of the informational subjects for even a week."¹ Standardized tests which might be used for such measurement are now available in most of the subjects of the elementary school curriculum and in some secondary subjects. Ordinarily they have been administered in September and June and comparisons have been made with norms at the beginning and end of the school year. Numerous questions arise that can not be answered by measurements made at such intervals. What are the amount and rate of progress throughout the year? Is the improvement, which is practically always found, uniform, or is the improvement rapid in the early part of the year with a gradual tapering off toward the close, or is the improvement slow at first and more rapid as the year advances? Are there plateaus in the progress curve? How do school subjects vary in improvability, both as to nature, amount, and rate of gain, and as to the

¹ Thorndike, E. L. *Educational psychology*. New York: Teachers College, 1913, Volume II: "Psychology of learning," p. 333.

factors and conditions which determine them? How do individuals differ in improvement? These and many other questions require more frequent and detailed records of progress than have hitherto been available.

The problem of studying improvement in school subjects throughout the year by means of standardized tests given at monthly intervals was proposed to Mr. L. J. Coubal, Principal of the Alice School, Hibbing, Minnesota, in 1915, as a thesis for a bachelor's degree. A portion of the material was later used by Mr. Edward Van Landegend for a study of individual differences in improvement. The results of both studies are so striking as to deserve wider publicity.

Tests were given to 32 pupils of the second grade, 33 of the third grade, 25 of the fourth grade, 18 of the fifth grade, and 20 of the sixth grade. Rate and comprehension in reading were measured by Starch's tests in the fourth grade only, as alternative tests could not be provided for all the other years. The scores for rate are in terms of the number of words read per second and for comprehension in terms of the number of words correctly reproduced. Rate of writing was determined for all the grades, and the quality was rated by the Ayres scale. Starch's spelling tests, consisting of one hundred words, were given to all the grades, the scores giving the number of words correctly spelled out of one hundred. Ability in addition, subtraction, multiplication, and division were measured by the Courtis Standard Research Tests, Series B, Forms 1, 2, and 3, in the fourth, fifth, and sixth grades. The tests, except in arithmetic, were given by the teachers under very specific directions, the principal also being present. The tests in arithmetic were all given by the principal of the school.

The results of the tests appear in Tables Ia to Ie inclusive.

TABLE Ia. RESULTS OF TESTS. GRADE II

| | Sept. 16 | Oct. 14 | Nov. 14 | Dec. 8 | Jan. 20 | Feb. 17 | Mar. 16 | Apr. 13 | May 11 | Gain thru Year | STANDARD SCORES FOR TESTS | |
|---------------|-------------|------------|------------|-----------|------------|------------|------------|------------|-----------|----------------------|---------------------------------|------|
| | | | | | | | | | | | Sept. | June |
| Writing: | | | | | | | | | | | | |
| Rate..... | | 10.9 | 10.4 | 27.5 | 24.8 | 33.8 | 25.9 | 31.9 | 33.1 | 22.2 | 20 | 31 |
| Quality..... | | 27.0 | 27.0 | 25.7 | 25.2 | 30.0 | 33.4 | 33.4 | 35.8 | 8.8 | | 27 |
| Spelling..... | 6.5 | 10.0 | 11.4 | 12.7 | 17.2 | 24.5 | 22.6 | 29.2 | 25.6 | 19.1 | 10 | 30 |

TABLE 1b. RESULTS OF TESTS. GRADE III

| | Sept. 16 | Oct. 14 | Nov. 11 | Dec. 8 | Jan. 20 | Feb. 17 | Mar. 16 | Apr. 13 | May 11 | Gain thru Year | STANDARD SCORES FOR TESTS | |
|---------------|-------------|------------|------------|-----------|------------|------------|------------|------------|-----------|----------------------|---------------------------------|------|
| | | | | | | | | | | | Sept. | June |
| Writing: | | | | | | | | | | | | |
| Rate..... | | 47.0 | 44.0 | 56.7 | 51.0 | 48.3 | 53.8 | 48.6 | 51.8 | 4.8 | 31 | 38 |
| Quality..... | | 37.0 | 39.0 | 43.0 | 28.4 | 39.2 | 33.3 | 35.0 | 32.4 | -4.6 | 27 | 33 |
| Spelling..... | 28.9 | 28.7 | 30.6 | 32.4 | 30.2 | 35.4 | 40.2 | 43.7 | 51.0 | 22.1 | 30 | 40 |

TABLE 1c. RESULTS OF TESTS. GRADE IV

| | Sept. 16 | Oct. 14 | Nov. 11 | Dec. 8 | Jan. 20 | Feb. 17 | Mar. 16 | Apr. 13 | May 11 | Gain thru Year | STANDARD SCORES FOR TESTS | |
|-------------------------|-------------|------------|------------|-----------|------------|------------|------------|------------|-----------|----------------------|---------------------------------|------|
| | | | | | | | | | | | Sept. | June |
| Reading: | | | | | | | | | | | | |
| Rate..... | 1.5 | 2.5 | 2.2 | 2.3 | 2.2 | 2.5 | 2.7 | 3.1 | 3.1 | 1.6 | 2.1 | 2.4 |
| Comprehension... | 31.0 | 23.2 | 32.7 | 36.3 | 36.4 | 40.9 | 40.1 | 49.1 | 40.4 | 9.4 | 24.0 | 28.0 |
| Writing: | | | | | | | | | | | | |
| Rate..... | | 49.5 | 53.1 | 49.9 | 57.7 | 51.3 | 65.2 | 52.6 | 54.7 | 5.2 | 38.0 | 47.0 |
| Quality..... | | 36.4 | 37.0 | 43.8 | 38.9 | 43.5 | 42.1 | 44.7 | 48.0 | 11.6 | 33.0 | 37.0 |
| Spelling..... | 41.1 | 49.2 | 39.0 | 48.7 | 45.7 | 49.6 | 55.4 | 56.1 | 56.7 | 15.6 | 40.0 | 51.0 |
| Arithmetic: | | | | | | | | | | | | |
| Addition { A..... | 5.7 | 5.0 | 6.4 | 7.8 | 7.4 | 8.5 | 10.4 | 12.5 | 11.5 | 5.8 | 4.0 | 6.0 |
| { R..... | 2.3 | 2.7 | 3.1 | 4.5 | 4.1 | 4.6 | 7.0 | 8.5 | 6.5 | 4.3 | 2.0 | 3.0 |
| Subtraction { A..... | 4.4 | 3.8 | 5.4 | 5.8 | 6.5 | 7.2 | 8.8 | 9.8 | 10.8 | 6.4 | 4.0 | 6.0 |
| { R..... | 1.4 | 1.4 | 3.0 | 3.7 | 4.5 | 3.3 | 6.5 | 8.0 | 8.0 | 6.6 | 1.0 | 3.0 |
| Multiplication { A..... | 3.5 | 3.5 | 3.7 | 5.0 | 5.6 | 6.0 | 7.5 | 8.4 | 9.1 | 5.6 | | 4.5 |
| { R..... | 1.1 | 1.8 | 1.3 | 2.8 | 3.5 | 3.0 | 5.0 | 5.5 | 6.4 | 5.3 | | 1.5 |
| Division { A..... | | | | 3.1 | 4.4 | 4.4 | 5.1 | 6.5 | 8.2 | 8.2 | | 3.5 |
| { R..... | | | | 0.9 | 0.9 | 0.9 | 2.1 | 5.0 | 6.2 | 6.2 | | 1.0 |

Complete results in reading, writing, spelling, and arithmetic were available for the fourth grade. In Figures 1-7 solid lines show graphically the progress for this grade from September to May. The standard September and June scores for these tests are shown by dashed lines. In these figures the letters along the base line refer to the months from September to May. No September scores are available for multiplication and division. Hence no dashed lines appear in Figures 6 and 7. June scores are indicated by a point at the right of each of these figures. Since the curves for the other grades show similar tendencies, the records of the fourth grade only will be used for this discussion.

TABLE Id. RESULTS OF TESTS. GRADE V

| | Sept. 16 | Oct. 14 | Nov. 14 | Dec. 8 | Jan. 20 | Feb. 17 | Mar. 16 | Apr. 13 | May 11 | Gain thru Year | STANDARD SCORES FOR TESTS | |
|-------------------------|-------------|------------|------------|-----------|------------|------------|------------|------------|-----------|----------------------|---------------------------------|------|
| | | | | | | | | | | | Sept. | June |
| Writing: | | | | | | | | | | | | |
| Rate..... | | 58.3 | 67.0 | 69.2 | 63.0 | 50.6 | 51.4 | 53.8 | 59.0 | 0.7 | 47.0 | 57.0 |
| Quality..... | | 37.6 | 45.6 | 47.5 | 47.1 | 46.0 | 49.3 | 48.0 | 52.3 | 14.7 | 37.0 | 43.0 |
| Spelling..... | 54.0 | 65.7 | 62.2 | 64.9 | 58.6 | 61.0 | 70.1 | 68.3 | 71.3 | 17.3 | 51.0 | 61.0 |
| Arithmetic: | | | | | | | | | | | | |
| Addition { A..... | 8.2 | 8.2 | 10.6 | 10.2 | 10.8 | 11.9 | 13.5 | 15.1 | 14.5 | 6.3 | 6.0 | 7.5 |
| { R..... | 5.2 | 5.0 | 6.9 | 7.0 | 7.1 | 8.1 | 11.2 | 12.5 | 10.2 | 5.0 | 3.0 | 4.0 |
| Subtraction { A..... | 6.9 | 5.8 | 8.7 | 8.8 | 10.1 | 11.5 | 14.6 | 14.2 | 16.2 | 9.3 | 6.0 | 8.0 |
| { R..... | 4.6 | 4.8 | 6.4 | 6.1 | 8.5 | 9.0 | 12.0 | 11.8 | 13.6 | 9.0 | 3.0 | 5.5 |
| Multiplication { A..... | 4.3 | 4.0 | 7.5 | 8.0 | 8.8 | 9.7 | 10.3 | 14.3 | 12.8 | 8.5 | 4.5 | 7.0 |
| { R..... | 1.8 | 1.7 | 4.9 | 4.3 | 5.4 | 6.3 | 7.6 | 11.5 | 9.6 | 7.8 | 1.5 | 4.0 |
| Division { A..... | 4.0 | 4.0 | 7.7 | 8.1 | 6.4 | 10.0 | 11.6 | 15.2 | 13.8 | 9.8 | 3.5 | 5.0 |
| { R..... | 1.4 | 1.9 | 5.2 | 5.6 | 4.1 | 7.5 | 10.1 | 14.1 | 12.3 | 10.9 | 1.0 | 3.0 |

TABLE Ie. RESULTS OF TESTS. GRADE VI

| | Sept. 16 | Oct. 14 | Nov. 14 | Dec. 8 | Jan. 20 | Feb. 17 | Mar. 16 | Apr. 13 | May 11 | Gain thru Year | STANDARD SCORES FOR TESTS | |
|-------------------------|-------------|------------|------------|-----------|------------|------------|------------|------------|-----------|----------------------|---------------------------------|------|
| | | | | | | | | | | | Sept. | June |
| Writing: | | | | | | | | | | | | |
| Rate..... | | 70.0 | 81.0 | 76.2 | 63.7 | 57.5 | 62.1 | 56.5 | 59.6 | —10.4 | 57.0 | 65.0 |
| Quality..... | | 46.7 | 53.7 | 55.5 | 54.3 | 55.8 | 57.2 | 57.3 | 61.1 | 14.4 | 43.0 | 53.0 |
| Spelling..... | 62.4 | 60.2 | 68.7 | 69.9 | 60.8 | 65.3 | 75.3 | 75.7 | 80.7 | 18.3 | 61.0 | 71.0 |
| Arithmetic: | | | | | | | | | | | | |
| Addition { A..... | 8.8 | 8.2 | 10.1 | 10.0 | 9.3 | 10.3 | 12.3 | 14.0 | 13.8 | 5.0 | 7.5 | 9.0 |
| { R..... | 5.6 | 4.8 | 7.1 | 6.3 | 6.2 | 7.0 | 10.0 | 11.5 | 10.3 | 4.7 | 4.0 | 5.0 |
| Subtraction { A..... | 6.5 | 7.3 | 8.7 | 8.8 | 8.8 | 10.0 | 12.1 | 12.9 | 13.4 | 6.9 | 8.0 | 10.0 |
| { R..... | 4.8 | 5.3 | 7.1 | 6.7 | 7.8 | 8.6 | 11.1 | 12.0 | 11.8 | 7.0 | 5.5 | 7.0 |
| Multiplication { A..... | 6.4 | 7.1 | 8.9 | 9.0 | 8.7 | 9.6 | 10.8 | 12.0 | 12.5 | 6.1 | 7.0 | 8.5 |
| { R..... | 4.6 | 5.2 | 6.9 | 5.6 | 7.4 | 7.9 | 9.1 | 10.0 | 9.4 | 4.8 | 4.0 | 5.5 |
| Division { A..... | 4.8 | 5.6 | 7.5 | 6.8 | 8.4 | 8.8 | 12.0 | 12.4 | 13.4 | 8.6 | 5.0 | 6.5 |
| { R..... | 3.8 | 4.3 | 6.4 | 5.9 | 7.5 | 7.4 | 10.4 | 11.0 | 12.9 | 9.1 | 3.0 | 5.0 |

The primary purposes of this study were (1) to measure as exactly as possible the amount of improvement throughout the year, and in particular to discover the effects of experimental conditions; (2) to find the typical curve of progress, if there is one, under such conditions; (3) to compare differences in rate of progress in different school subjects; and (4) to secure data on the

effects of equal practice on individual differences. It was thought that the giving of the tests at regular intervals might be a valuable incentive and moreover would be of great use to a supervisor in diagnosing defects in individuals and detecting inequalities in emphasis or other irregularities.

FIG. 1.
READING.

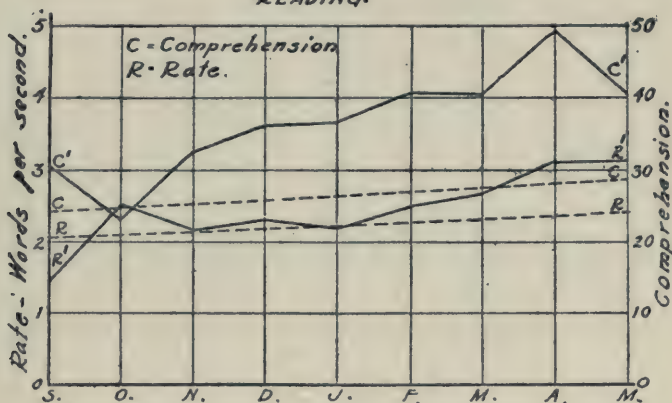
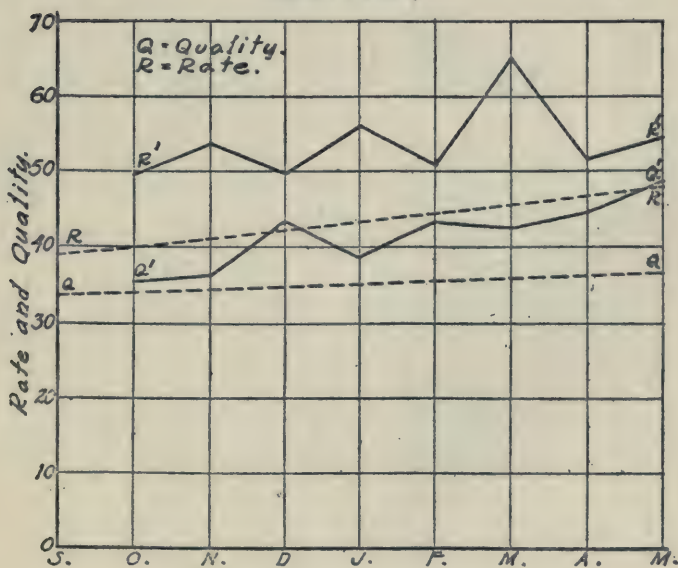
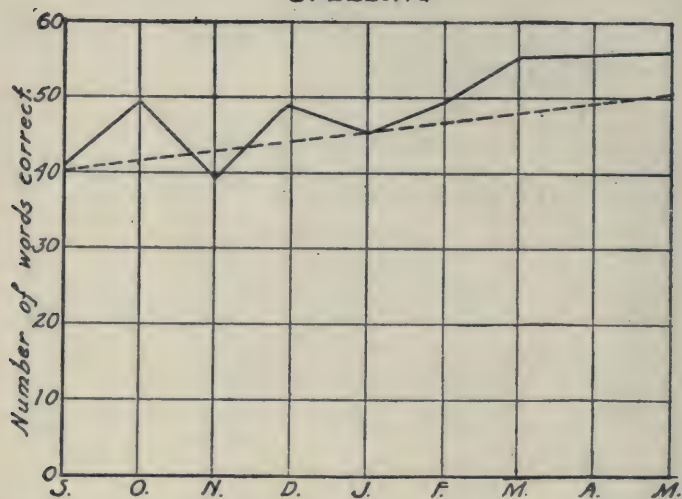


FIG. 2.
WRITING.



•FIG. 3.
•SPELLING.



•FIG. 4.
•ADDITION.

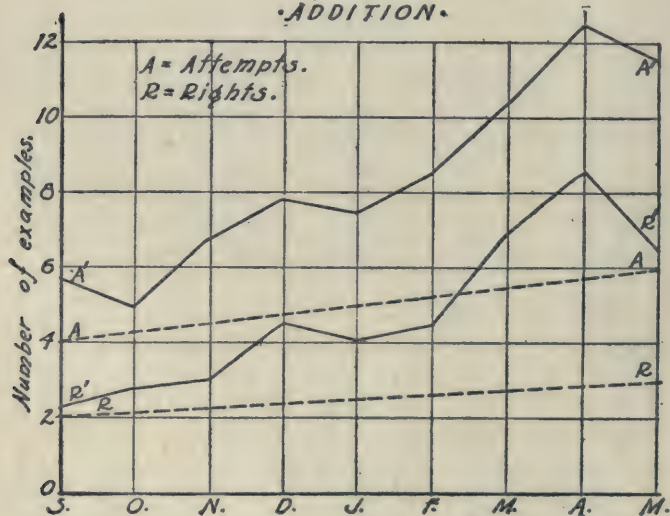


FIG 5
· SUBTRACTION ·

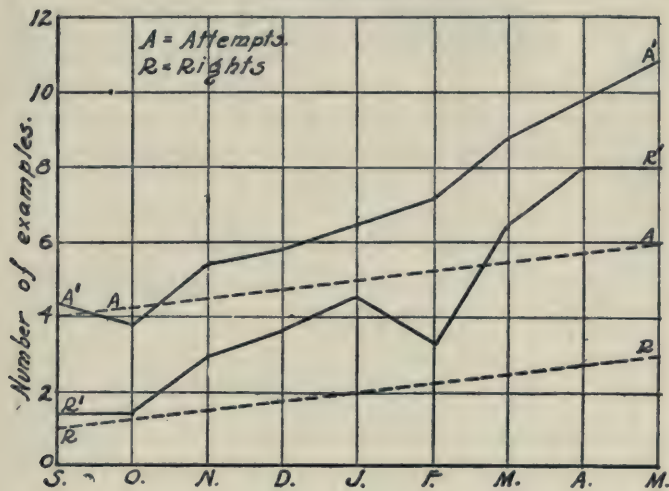


FIG 6
· MULTIPLICATION ·

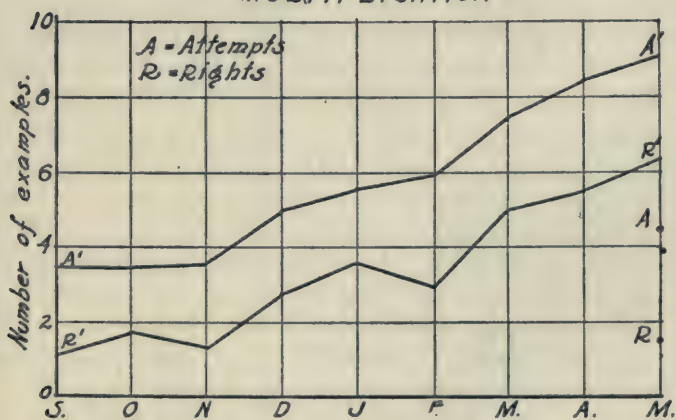
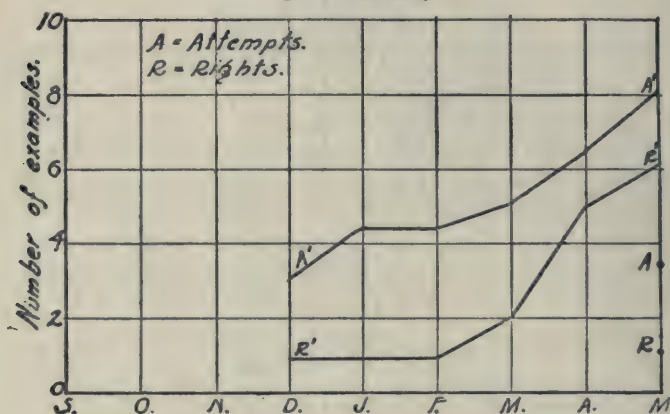


FIG. 7.
DIVISION.



The figures and tables bring out certain significant facts:

1. *The astonishingly great amount of improvement.*—The effect of applying the tests at regular intervals throughout the year is nothing short of remarkable. This fourth grade group of 25 pupils began the year with scores approximating closely the standards. At the end of the year the gains exceeded all expectations. In rate of reading the initial score of 1.5 words per second in September is considerably below the standard, while the May score of 3.1 words attains the standard of the sixth grade. In comprehension the initial score of 31 words reproduced is above the standard considerably, rate having been sacrificed for comprehension in the first month, while the May score of 40.4 words is again that of the sixth grade. In spelling the September score of 41.1 words is very close to the standard, while the May score of 56.7 words is almost that of the fifth grade. In writing the initial scores approximate closely the standard June scores for the grade, while the final score for rate is that of the fifth grade and for quality midway between those of the fifth and sixth grades. In the four fundamental operations the gains are even more striking. The number of rights runs in addition from 2.3 to 6.5 (standards 2.0 to 3.0), in subtraction from 1.4 to 8.0 (standards 1.0 to 3.0), in multiplication from 1.1 to 6.4 (June standard 1.5) and in division the final score is 6.2 (standard 1.0). The final scores in

arithmetic in attempts and rights are in all cases above the standards of the sixth grade and in many cases are above those of the seventh grade. When it is borne in mind that selection by eliminating the poorer pupils accounts in some measure for the increase in the standard scores from year to year, the effect is all the more noteworthy.

It should be noted also that the gain in one subject is not due to overemphasis upon it and accounted for in part by loss in other subjects. The results of a very large share of the instruction of the fourth grade were measured and large gains are shown in every case. Nor is the increase due to constant drill. When a test was completed it was put aside and not used again till the end of the next month. No alteration was made in the work of the school except for the monthly tests. The regular work of the fourth grade involves a five to ten minute daily exercise in the four fundamental operations in arithmetic.

The results are in accord with those attained in laboratories which have always shown very great improvement with practice under experimental conditions even in traits which in the ordinary circumstances of life have been much subject to practice. Laboratory experiments in memorizing, for instance, have shown very great gains under experimental conditions by those who have memorized more or less all their lives and whose capacities might be supposed to have been brought near the limits of improvement. In sharp contrast is the rate of improvement ordinarily found in school children as is indicated by the relatively small gain from September to June in the standard scores.

Certain practical applications of these results suggest themselves. Would it not be desirable to place a great deal of school work in the form of a practice experiment in which the pupil tested himself at regular intervals by objective methods, kept his own score, and watched his gain or loss from week to week or month to month? All experiments show that this would be an extraordinary stimulus to effort. The individual would be competing not with others but with his own past performances. Why should not a pupil record at intervals how fast he reads, how much he comprehends and recalls, how well he writes and spells, where he stands in composition, how many words he knows in a foreign language, how rapidly and accurately he can

read or translate, how long he must study a lesson assignment, etc.? The results of these experiments in one school show that the experimental conditions subjected the school to a healthy strain and that an amount of improvement resulted greater than any that has to my knowledge been reported.

2. *The rate of improvement.*—An inspection of the figures for the fourth grade as well as for the other grades shows that progress throughout the year is not one of uniform growth. No single curve or generalization describes all of the facts, but a curve with a long slow rise during the first half or two-thirds of a year, with a rapid rise during the last third, fits the facts tolerably well. This curve differs markedly from the typical learning curve which shows a rapid rise at first, followed by a period of no gain or of slow gain. Comparisons with the typical learning curve, where relatively new coordinations are set up, can, of course, not be fairly made since these curves of progress in school subjects are but segments of larger curves. Whether the typical curve of progress in school subjects is the one found so frequently in these records, it will require further experimentation to determine. It will also require further experimentation to determine how far the long plateau period in the early part of the year is necessary. It is quite possible that changes in methods and devices, or furnishing other stimulations, might bring about greater improvement earlier in the year. At any rate the facts stand out pretty sharply, challenge one's interest, and invite to further study.

3. *The differences in rate of improvement in the different subjects.*—The least improvement is in spelling ability; the greatest in arithmetic, particularly in addition, both in rate and accuracy. These results agree with those of Kirby.² Arithmetical operations reduce to a definite number of combinations, which recur frequently, are easily isolated and made the subject of training; the associations required in spelling occur less frequently and would show a less rapid improvement. We are just at the beginning of the study of such differences as these in school subjects, the types of learning they represent, and the factors and conditions of improvement. The nature and rate of improvement in history and geography differ no doubt in important respects

² Kirby, T. J. *Practice in case of school children*. New York: Teachers College, 1913.

from those in reading, writing, spelling, and arithmetic. There is great need for study of learning processes in each school subject under school conditions, using standardized tests as instruments of analysis. Such tests have furnished a valuable means of

TABLE II. INDIVIDUAL DIFFERENCES IN ADDITION, SUBTRACTION, AND MULTIPLICATION COMBINED. GRADE IV

Scores=Sum of Rights in Courtis Standard Research Tests in Arithmetic, Series B

| Pupil | S | O | N | D | J | F | M | A | M | Gross Gain |
|---------------------|-----|------|------|------|------|------|------|------|------|------------|
| 1..... | 13 | 15 | 13 | 18 | | 23 | 31 | 35 | 38 | 25 |
| 2..... | 1 | 6 | 9 | 18 | 17 | 14 | 31 | 34 | 37 | 36 |
| 3..... | 13 | 16 | 15 | 25 | 22 | 20 | 31 | 41 | 32 | 19 |
| 4..... | 11 | 8 | 10 | 19 | 21 | | 23 | 34 | 31 | 20 |
| 5..... | 5 | 6 | 12 | 16 | 18 | 18 | 25 | 29 | 28 | 23 |
| Av. Group I..... | 8.6 | 10.2 | 11.8 | 19.2 | 19.5 | 18.8 | 28.2 | 34.6 | 33.2 | 24.6 |
| 6..... | 2 | 5 | 12 | 11 | 11 | 12 | 21 | 20 | 28 | 26 |
| 7..... | 4 | 4 | 11 | 15 | 15 | 12 | 20 | 28 | 24 | 20 |
| 8..... | 4 | 5 | 8 | 7 | 12 | 12 | 25 | 26 | 23 | 19 |
| 9..... | 9 | 7 | 14 | 15 | 13 | 15 | 20 | 23 | 22 | 13 |
| 10..... | 5 | 3 | 9 | 10 | 14 | 14 | 18 | 21 | 20 | 15 |
| Av. Group II..... | 4.8 | 4.8 | 10.8 | 11.6 | 13.0 | 13.0 | 20.8 | 23.6 | 23.4 | 18.6 |
| 11..... | 9 | 13 | 12 | 17 | 13 | 8 | 21 | 22 | 18 | 9 |
| 12..... | 7 | 7 | 6 | 8 | 5 | 7 | 16 | 16 | 17 | 10 |
| 13..... | 3 | 6 | 3 | 7 | 8 | 5 | 14 | 19 | 17 | 14 |
| 14..... | 1 | 2 | 4 | 8 | 10 | 9 | 14 | 21 | 17 | 16 |
| 15..... | 6 | 10 | 5 | 7 | | 12 | 18 | 6 | 16 | 10 |
| Av. Group III..... | 5.2 | 7.4 | 6.0 | 9.4 | 9.0 | 8.2 | 16.6 | 16.8 | 17.0 | 11.8 |
| 16..... | 0 | 0 | 3 | 4 | 12 | 7 | 10 | 13 | 16 | 16 |
| 17..... | 0 | 4 | 5 | 6 | 2 | 10 | 5 | 8 | 10 | 10 |
| 18..... | 0 | 7 | 5 | 4 | 9 | 6 | | 11 | 9 | 9 |
| 19..... | 3 | 2 | | 11 | 12 | 4 | 9 | | 9 | 6 |
| 20..... | 4 | 3 | 2 | 0 | 4 | 0 | 0 | 11 | 6 | 2 |
| Av. Group IV..... | 1.4 | 3.2 | 3.8 | 5.0 | 7.4 | 5.4 | 6.0 | 10.8 | 10.0 | 8.6 |
| Class Averages..... | 5.0 | 6.4 | 8.3 | 11.3 | 12.1 | 10.9 | 18.5 | 22.0 | 20.9 | 15.9 |

diagnosis, but progress in therapeutics has not kept pace with progress in diagnosis.

4. *Individual differences in amount and rate of improvement.*—To what are the great gains with practice due? Are they due to a uniform gain by all pupils, are they due to very great gains by the better pupils, or are they due to large gains by the poorer pupils? Do the initial individual differences tend to increase or decrease throughout the year? The few studies that have been made have shown marked differences in improvement and, moreover, they have shown that individual differences tend to increase rather than to decrease with equal amounts of practice.³ The data in this study point unmistakably to the same conclusion. A

TABLE III. INDIVIDUAL DIFFERENCES IN FOUR FUNDAMENTAL OPERATIONS. GRADE V

Scores=Sum of Rights in Courtis Standard Research Tests in Arithmetic, Series B

| Pupil | S | O | N | D | J | F | M | A | M | Gross Gain |
|--------------------|------|------|------|------|------|------|------|------|------|------------|
| 1..... | 9 | 12 | 21 | 36 | 36 | 47 | 64 | 65 | 64 | 55 |
| 2..... | 28 | 25 | 31 | 39 | 40 | 44 | 50 | 63 | 63 | 35 |
| 3..... | 8 | 20 | 20 | 17 | 19 | 28 | 47 | 69 | 63 | 55 |
| 4..... | 18 | 20 | 31 | 34 | 30 | 39 | 47 | 43 | 56 | 38 |
| Av. Group I..... | 15.8 | 18.8 | 25.8 | 31.5 | 31.2 | 39.5 | 52.0 | 60.0 | 61.5 | 45.7 |
| 5..... | 17 | 17 | 33 | 33 | 33 | 48 | 56 | 62 | 51 | 34 |
| 6..... | 11 | 11 | 22 | 22 | 23 | 27 | 33 | 46 | 44 | 33 |
| 7..... | 19 | 16 | 29 | 34 | 20 | 38 | 40 | 42 | 42 | 23 |
| 8..... | 6 | 5 | 21 | 21 | 17 | 20 | 31 | 53 | 41 | 35 |
| Av. Group II..... | 13.3 | 12.3 | 26.3 | 27.5 | 23.3 | 33.3 | 40.0 | 50.8 | 44.5 | 31.2 |
| 9..... | 18 | 16 | 21 | | 23 | 28 | 38 | 52 | 39 | 21 |
| 10..... | 8 | 17 | 20 | 22 | 25 | 20 | 31 | 37 | 38 | 30 |
| 11..... | 13 | 14 | 15 | 19 | | 25 | 30 | | 33 | 20 |
| 12..... | 7 | 7 | 11 | 11 | 18 | 11 | 22 | 19 | 20 | 13 |
| Av. Group III..... | 11.5 | 13.5 | 16.8 | 17.3 | 22.0 | 21.0 | 30.3 | 36.0 | 32.5 | 21.0 |
| Class Average..... | 13.5 | 15.0 | 22.9 | 26.2 | 25.8 | 31.3 | 40.7 | 50.1 | 46.2 | 32.7 |

³ Thorndike, E. L., *Educational psychology, briefer course*. New York: Teachers College, 1914. pp. 390-94.

special study was made of individual differences in arithmetic in the fourth, fifth, and sixth grades. For this purpose a single score for each pupil was secured by adding together the number of examples right in each of the four fundamental operations. In the fourth grade the score is the sum of the rights in addition, subtraction, and multiplication, as the work in division was not begun until December. The twenty pupils of the fourth grade were divided into quartiles on the basis of the May score. The pupils of the fifth and sixth grade into tertiles.

Tables II-IV give the individual group and class scores for each month and the gross gains through the year.

TABLE IV. INDIVIDUAL DIFFERENCES IN FOUR FUNDAMENTAL OPERATIONS. GRADE VI

Scores=Sum of Rights in Courtis Standard Research Tests in Arithmetic, Series B

| Pupil | S | O | N | D | J | F | M | A | M | Gross Gain |
|--------------------|------|------|------|------|------|------|------|------|------|------------|
| 1..... | 28 | 38 | 48 | 56 | 50 | 57 | 69 | 66 | 79 | 51 |
| 2..... | 30 | 30 | 38 | 45 | 42 | 48 | 61 | 67 | 62 | 32 |
| 3..... | 21 | 20 | 36 | 28 | | 38 | 51 | 65 | 61 | 40 |
| 4..... | 21 | 19 | 31 | 24 | 33 | 35 | 43 | 56 | 61 | 40 |
| 5..... | 23 | 21 | 29 | 27 | 30 | 35 | 46 | 53 | 52 | 29 |
| Av. Group I..... | 24.6 | 25.6 | 36.4 | 36.0 | 38.8 | 42.6 | 54.0 | 61.4 | 63.0 | 38.4 |
| 6..... | 34 | 40 | 49 | 32 | 36 | 34 | 50 | 58 | 49 | 15 |
| 7..... | 16 | 19 | 36 | 34 | 33 | 36 | 48 | 48 | 45 | 29 |
| 8..... | | 20 | 28 | 30 | 28 | 30 | 41 | 45 | 42 | 22 |
| 9..... | 14 | 21 | 23 | 27 | 27 | 27 | 43 | 35 | 39 | 25 |
| 10..... | 9 | 6 | 16 | 12 | 21 | 19 | 27 | 37 | 39 | 30 |
| Av. Group II..... | 18.3 | 21.2 | 30.4 | 27.0 | 29.0 | 29.2 | 41.8 | 44.6 | 42.8 | 24.2 |
| 11..... | 8 | 15 | 13 | 12 | | 21 | 35 | 36 | 38 | 30 |
| 12..... | | 16 | 18 | 18 | 24 | 30 | 34 | 35 | 37 | 21 |
| 13..... | 12 | 10 | 17 | 12 | | 16 | 30 | 28 | 35 | 23 |
| 14..... | 17 | 24 | 30 | 24 | 26 | | 25 | 28 | 30 | 13 |
| 15..... | | 13 | 20 | 12 | 22 | 17 | 27 | 31 | 23 | 10 |
| Av. Group III..... | 12.3 | 19.5 | 19.6 | 15.6 | 24.0 | 21.0 | 30.2 | 31.6 | 32.6 | 19.2 |
| Class Average..... | 19.4 | 20.8 | 28.8 | 26.2 | 31.0 | 31.5 | 42.0 | 45.9 | 46.1 | 26.7 |

The problem of measuring accurately the amount of improvement with practice, and especially the comparison of individuals with varying initial and final efficiency scores in amount and rate of gain, is beset with difficulties, as has been frequently pointed out.⁴ While it is therefore impossible to state adequately or definitely the amounts and rates of gain by either a gross gain or a percentile method, an inspection of the tables shows that by any rational method of interpretation the scriptural law "to him that hath shall be given" holds. Those with the highest initial scores have the highest final scores and gain most both absolutely and relatively. This is perhaps best indicated by expressing the scores in terms of deviations from the averages. Tables Va-Vc inclusive give these facts for the three grades.

TABLE Va. GROUP DEVIATIONS FROM CLASS AVERAGES. GRADE IV

| | S | O | N | D | J | F | M | A | M |
|--------------|------|------|------|------|------|------|-------|-------|-------|
| Group I..... | +3.6 | +3.8 | +3.5 | +8.9 | +7.4 | +7.9 | +9.7 | +12.6 | +12.3 |
| " II..... | -0.2 | -1.6 | +2.5 | +0.3 | +1.1 | +2.1 | +2.3 | +1.6 | +2.5 |
| " III..... | +0.2 | +1.0 | -1.7 | -1.9 | -3.1 | -2.7 | -1.9 | -5.2 | -3.9 |
| " IV..... | -3.6 | -3.2 | -4.5 | -6.3 | -4.7 | -5.5 | -12.5 | -11.2 | -10.9 |

TABLE Vb. GROUP DEVIATIONS FROM CLASS AVERAGES. GRADE V

| | S | O | N | D | J | F | M | A | M |
|--------------|------|------|------|------|------|-------|-------|-------|-------|
| Group I..... | +2.3 | +3.8 | +2.9 | +5.3 | +5.4 | +8.2 | +11.3 | +9.9 | +15.3 |
| " II..... | -0.2 | -2.7 | +3.4 | +1.3 | -2.5 | +2.0 | -0.7 | +0.7 | -1.7 |
| " III..... | -2.0 | -1.5 | -6.1 | -8.9 | -3.8 | -10.3 | -10.4 | -14.1 | -13.7 |

TABLE Vc. GROUP DEVIATIONS FROM CLASS AVERAGES. GRADE VI

| | S | O | N | D | J | F | M | A | M |
|--------------|------|------|------|-------|------|-------|-------|-------|-------|
| Group I..... | +5.2 | +4.8 | +7.6 | +9.8 | +7.8 | +11.1 | +12.0 | +15.5 | +16.9 |
| " II..... | -1.1 | +0.4 | +1.6 | +0.8 | -0.2 | -2.3 | -0.2 | -1.3 | -3.3 |
| " III..... | -7.1 | -1.3 | -9.2 | -10.6 | -7.0 | -10.5 | -11.8 | -14.3 | -13.5 |

⁴ Thorndike, E. L., *Psychology of learning*, 1913, pp. 165-77.

It is apparent that Group I in each case deviates farther and farther above the average with each succeeding month, while Group IV in the fourth and Groups III in the fifth and sixth grades drop farther and farther below the average. The highest group progresses with a marked positive acceleration, while the lowest group progresses with equally marked negative acceleration in relation to the class average. The result is that while Group IV in the fourth grade attains the standard score of the grade, Group I with an initial score of the fourth grade in June attains almost that of a standard eighth grade. The range of variation in progress in the fifth and sixth grades is similarly great—far beyond the range of standards published.

This study is but a beginning in an important problem, *viz.*, to record more accurately than we have hitherto the progress of children through the school year. Progress is determined by a great number of factors and conditions, all of which need to be carefully studied. The great effect of placing work under experimental conditions as a stimulus to interest and effort is amply demonstrated. Its value for supervisory purposes in diagnosis and classification is equally clear.

IMPROVEMENT OF INSTRUCTION THROUGH THE USE OF EDUCATIONAL TESTS

WALTER S. MONROE

Assistant Director Bureau of Educational Research, University of Illinois

A few years ago in conversation with the superintendent of schools in a city of about forty thousand population, I inquired concerning the use he was making of educational tests; he replied that he was using tests, that in fact, he made it a point to give each new test to a few hundred pupils. He then made a very significant statement. He said that he had given the Courtis arithmetic tests the year before, and that his teachers had failed to find the pupils doing any better in arithmetic after the tests were given than they did before.

Unfortunately, this incident represents the attitude of a considerable number of users of standardized tests. They consider such tests as teaching devices, and express dissatisfaction with them when they do not cause pupils to do better school work. Others have thought of educational tests as playthings which they might use for their own entertainment or for the entertainment of audiences at teachers' meetings. Beautiful charts in colored inks have been constructed to represent the scores obtained; and because such charts are new and tangible, they have attracted much attention.

Educational tests are not playthings; neither are they teaching devices. It is not their function to increase the abilities of pupils. They are designed to yield information in the form of measures of the abilities of pupils. Only as this information is interpreted in terms of pupil needs and these are in turn used as a basis in planning future instruction, will the use of tests increase abilities.

In the past the major use of educational tests has been by superintendents and principals. Their purpose has been to ascertain how the pupils of their schools stood in comparison with pupils in corresponding grades of other schools or with reference to standard scores. In some cases the tests have been used as a

basis for judging the efficiency of teachers. These uses are important and in the last analysis have as their function the improvement of instruction. Educational tests, however, may be used and indeed ought to be used directly by teachers as instruments for making more effective their efforts as instructors. We now have sufficient scientific evidence to show that the measurement of the abilities of pupils at appropriate intervals and the use of this information in planning future teaching will increase the abilities of the pupils. Hence, it is important that teachers learn how to use standardized tests to improve instruction.

Giving these tests, marking the papers, and tabulating the scores are only first steps. In many instances beneficial results have followed when the work has not been carried beyond this stage. The attention of both pupils and teachers may thus have been directed to the minimum essentials of the school subject. They may have been given a more definite aim and a stronger motive. But, in order to reap the full benefit, the scores obtained from giving a test must be interpreted in terms of pupil needs; and then the instruction must be modified in accordance with the interpretation. This article is, therefore, written to suggest a method of interpreting the scores of a group of pupils.

Interpretation in terms of pupil needs involves more than simply ascertaining that the class median is above, up to, or below standard or that certain pupils are above, up to, or below standard. Doing this is equivalent to a physician's telling his patient that his temperature is 103, that his liver is not functioning properly, that he has serious indigestion, and that he is threatened with a nervous breakdown. Clearly this is insufficient. If the physician is to be of service to his patient, he must interpret these facts in terms of his patient's needs. Then in accordance with these needs, the physician must prescribe for him—for example, tell him to eat no meat or starchy foods, to spend several hours out of doors each day, and to take certain medicines before each meal and at bedtime.

In interpreting scores of educational tests in terms of pupil needs, it is necessary to make similar prescriptions of changes in instruction. The scores of the pupils and of the class, together with the errors, correspond to the temperature, pulse rate, condition of tongue, location of pain, etc., of a patient. They are

the *symptoms*; and it is their meanings which are of real importance. Besides saying that the median of a certain sixth grade class in silent reading is below the standard for the fifth grade and that 25 percent of the pupils are below standard for the fourth grade, while only 18 percent are up to or above standard, we should add that as a whole the class needs to give less attention to oral reading, that it needs to devote a practice period of 30 minutes a day to the silent reading of simple descriptive prose, and that certain specified pupils need to give special attention to acquiring the meaning of words. An interpretation may also include a statement that certain pupils who made very high scores ought to be promoted to the next grade, or that they need not continue to study silent reading, but that they should spend the time thus saved in the study of arithmetic.

Different classes as a whole have different needs, and the individual members of a class have a great variety of pupil needs. The first step, however, in interpreting the scores of a class is to ascertain the needs that are common to the class as a whole or to large groups within the class. This should be followed by an individual interpretation of the scores of those pupils at the extremes of the group, particularly of those conspicuously below standard. At present, we do not have sufficient scientific information to permit us to formulate complete statements of the pupil needs which are indicated by the various combinations of symptoms. In the absence of such formulations, it will be helpful to outline a general procedure which may be followed in interpreting the scores of a class after they have been tabulated on the class record sheet. In order that the subject may be as concrete as possible, it will be stated in terms of the Courtis Standard Research Tests, Series B. However, one can easily adapt the general plan to any other test.

In a rough way we may group the great variety of combinations of scores of classes under six cases as follows:

- Case 1. Class medians below standard in both rate and accuracy.
- Case 2. Class medians for rate below standard with satisfactory or high accuracy medians.
- Case 3. Class medians for accuracy below standard with satisfactory or high rate medians.
- Case 4. Scores of members of the class widely scattered.

Case 5. Irregular development; class medians in one test up to or above standard with class medians in another test distinctly below standard.

Case 6. Class medians for both rate and accuracy up to or above standard with the individual scores grouped closely about the medians.

Any particular class may exhibit certain combinations of these conditions. For example, the median rate in a test (i.e., addition) may be below standard with a satisfactory or high accuracy median and at the same time the scores of the members of the class may be widely scattered. Such a class would be considered as coming under both Case 2 and Case 4.

For each of these cases, it is possible to list certain probable needs. All of these needs may not apply to a given class but some of them will almost certainly be appropriate. In general, it will be profitable to investigate these needs first and determine which of them apply to the class in question. The most probable needs for each case are given below.

Case 1.—This is probably a case of inefficient teaching. It is true, there may be some other explanation of the low median scores and it is well of course to investigate this possibility. For example, it may be suspected that pupils are below normal in general ability, and evidence on this question may be obtained by giving a test of general intelligence. Or, scores may be below standard because of an unusual course of study, and a comparison of the particular course with a number of successful courses may be made. Still other causes may be responsible to a greater or less extent for the condition; but inefficient teaching is the most probable single cause.

Instruction may be inefficient with respect to drill. Perhaps more time should be devoted to it, but it is more likely that the time already devoted to it should be used more effectively. Drill work ought to be organized so that every minute is used by every pupil. Many teachers waste time by allowing some pupils to be idle while others are occupied. When pupils are working examples at the board, some teachers dictate one example at a time and do not stop the work until the slowest pupil has finished. The result is that pupils who work rapidly are idle a large part of the time. It is easy to carry on drill work in such a way that loss of this kind will not occur. In the case just mentioned the

amount of time wasted will be greatly reduced if several examples instead of one are dictated at a time and the work stopped as soon as the most rapid workers have finished. Some teachers also waste time in having purely mechanical work explained when the pupils need drill rather than explanation. Again, drill may be inefficient because of the failure of the teacher to recognize *rate* of work as important. The need may be increased motivation. If this is the case, the teacher will find educational tests themselves helpful. In addition he can make use of a variety of devices for securing a stronger motive, including the assignment of work under timed conditions.

Case 2.—The most probable need of a class which is below standard in rate with satisfactory accuracy is for *speed drills*. In other words, it is likely that the teacher has not been placing sufficient emphasis upon the rate of work. A possible need is for the elimination of time-consuming methods of work, such as the use of an elaborate phraseology in performing the operations. For example, pupils may be required to name each successive digit in adding a column of figures instead of naming only the partial sums. In some cases slow work is due to the failure of pupils to concentrate attention upon the task. Under such circumstances they should be trained in giving continuous and undivided attention to the task assigned. This includes, among other things, the need for securing a stronger motive.

Case 3.—A low median score in accuracy may be due to the fact that, in working under the timed conditions of the test, the pupils become excited and that in attempting to do a large number of examples they become careless. When this happens, the test should be repeated; and it is likely that on the second trial, the median for accuracy will be much higher. If, however, the condition persists, the pupils should have frequent timed drills under the conditions of the test and should thus learn to assume the proper attitude toward the test. It may be that the need is even more fundamental. There are many different types of examples in a given operation—e.g., in addition. Pupils may have learned to do some types without having learned to do others. For example, they may be able to add short columns without being able to add long columns because of the increased span of attention required. It may be that the pupils have never been given instruction upon the particular type of example

occurring in the test. If so, they need to be taught how to do this type.

Case 4.—When the scores of the several members of the class are too widely scattered, the probable need is for individual instruction. The form of this instruction will depend upon the particular needs of the individual pupils. An analysis of the performance of a pupil is required in order to know his particular needs; but in general, the probable needs of individual pupils are the same as the probable needs of a class. The conditions occurring in Case 4 may often be met by promoting certain pupils to a higher grade and demoting others to a lower grade, provided the general organization of the school will permit such action. When this cannot be done, the teacher must devise some way to give to each pupil that instruction which will meet his needs. Practice exercises such as the Courtis Standard Practice Tests¹ have been designed to assist in meeting individual needs.

Case 5.—Irregular development is made possible by the fact that there are a number of different types of examples. Pupils may and do learn to do certain types without learning to do other types with anything like the same rate and accuracy. When irregular development is found to exist, the need is for a redistribution of emphasis in the instruction. Clearly, types for which the scores are low need to be given more emphasis while those for which the scores are up to or above standard need less attention. The Courtis Standard Research Tests, Series B, can only show irregularities between the four fundamental operations. A series of diagnostic tests will reveal irregularities between different types of examples within the same operation.

Case 6.—When a class is above standard in all respects it may be thought that conditions are satisfactory and that the pupils have no needs. This, however, may not be true. Among the most probable needs are (1) promotion to the next grade, (2) less time to this subject and more to others, (3) opportunity to take up more advanced topics. The meeting of these needs (especially

¹ The Courtis Standard Practice Tests in Arithmetic (published by the World Book Company, Yonkers-on-Hudson, New York) are practice tests in the four fundamental operations with whole numbers. The complete set consists of cabinet 1 (a set of 720 lesson cards for a class of fifty pupils), cabinet 2 (a set of 288 lesson cards for a class of twenty), a student's record and practice pad, a teacher's manual, and a teacher's record. A specimen set may be obtained from the publishers for a dollar.

the one involving promotion to the next grade) depends on the general plan of organization of the school. When a class is just up to standard, pupils may need a continuation of the present instruction or they may now be ready to go on to something else.

Locating a class under one of these cases and ascertaining which of the suggested probable needs apply are the first steps in interpretation. With the partial exception of Case 4 (irregular distribution of scores), it is mass interpretation and for that reason it is necessarily crude. Like mass instruction, it will not fit all pupils. While the members of the class probably have some common needs, there are also likely to be many individual needs. Thus such mass interpretation should be supplemented by an interpretation of the scores of individual pupils, especially of those having the highest and lowest scores.

In both class interpretation and individual interpretation, it is necessary to remember that the teacher has access not only to the scores but also to the test papers which give the performances of the pupils in detail. An analysis of these is helpful, especially where there is any uncertainty about the needs of the pupils. A score reveals nothing as to what was lacking in the mental processes of the pupil, but an analysis of his errors may reveal the defect and suggest corrective instruction. In making this analysis, it will be helpful to classify the errors which the pupil has made. In multiplication, the error may be one of multiplication, of placing the partial products, or of addition. In addition, and to a less extent in the other operations, it is frequently impossible to determine the nature of the error. In such cases, one should observe the pupil as he works or even have him "express orally his mental processes." Whenever such detailed studies of the performances of pupils have been made, significant facts have been revealed. This kind of study requires much time, but the time is profitably employed.

Although interpretation cannot be reduced to a mechanical basis, it is believed that the general procedure just outlined will be helpful as a beginning and will lead naturally to the individual analyses without which accurate individual adjustments cannot be made. The fundamental thing for the teacher to bear in mind is the principle that only when we interpret the scores of a test in terms of pupil needs and modify instruction to meet those needs will educational tests fulfill their function.

VALUE OF INFORMAL TESTS OF READING ACCOMPLISHMENT

WILLIAM S. GRAY

School of Education, University of Chicago

Informal tests, as the term is used in this discussion, relate to tests which are organized by the classroom teacher or supervisor for the purpose of securing accurate records concerning the accomplishments of pupils. They differ from standardized tests in that they are not so carefully organized, they have not been given to a large number of pupils under similar conditions, and there are no standards which can be used as a basis for comparison. It is evident at once that such tests have serious limitations as contrasted with the more elaborate types of standardized tests. In recommending their use in this discussion it is not intended to urge that they take the place of standardized tests. It is intended on the other hand to show that informal tests must be used frequently along with standardized tests, if testing is to be continuous and most effective.

Directions for an informal test of oral reading follow in order to illustrate the character of such tests:

"Measuring the rate of oral reading.—The simplest measurement under the ordinary conditions of class work is probably a measurement of the rate of oral reading. This record can be secured in any ordinary reading exercise as follows: Choose for this exercise one or more selections which the pupils read early in the year. Set aside a reading period for a great deal of oral reading. Make no previous assignment. The pupils should read as in ordinary reading lessons without suspecting that their speed of reading is being recorded. Have at hand a sufficient quantity of reading material so that no two pupils read the same paragraphs.

"The teacher will need a reader which can be marked freely with a pencil; also a watch with a second hand. When a pupil begins reading, note the exact second at which he reads the first word of the selection. Follow the watch carefully, and when

exactly sixty seconds have passed, check the word which the pupil was reading when the minute was up. Let the pupil continue reading without interruption to the end of the paragraph as he would in any reading exercise. Make whatever comments or ask whatever questions would be asked in an ordinary reading lesson and then go on to the next pupil. If, during the reading, a pupil hesitates a few seconds because of a difficult word, pronounce the word for him so that he may continue reading.

"Be sure that you mark the following points for each member of the class: (a) the particular part read by each pupil; (b) the point at which he began reading; (c) the point which had been reached at the end of sixty seconds. At your earliest convenience count the number of words read by each pupil during the minute. Count "a" and "the" and other short words, each as one word. Record the total number opposite the pupil's name on a record sheet which you will keep for your own use. Do not destroy these records. Keep them on file. They will be valuable for purposes of future comparison.

"This type of exercise should be continued for two or three class periods. Usually a teacher who tries to give tests and record facts for the first time finds that she is distracted and inaccurate. Confidence and skill develop very rapidly, however, with practice. Continue the measurements of the rate of oral reading until you feel confident that you can secure the records accurately."

There are three lines of argument which can be advanced in favor of informal tests. In the first place, the number of standardized tests in reading is limited. At most there are not more than ten which have been carefully standardized. Although these tests measure the accomplishments of pupils in several phases of reading ability they do not include, by any means, all of the significant phases. Until a more inclusive set of reading tests has been developed, teachers and supervisors will have to organize informal tests for their own use, if detailed studies of all phases of reading are made. Furthermore, only a limited number of tests appear in duplicate or triplicate form. Hence only a few standardized tests which are comparable in difficulty can be given at different times during the year. It is evident, therefore, that if teachers and supervisors carry on investigations of reading con-

tinuously, informal tests must be used at frequent intervals to supplement the more detailed studies made through the use of standardized tests. Unless investigations proceed regularly and systematically throughout the year, it is probable that tests will be interpreted as a novelty. Under these conditions they fail to enlist the constructive effort of teachers for any significant period of time.

In the second place, the wide use of informal tests brings to the teacher's attention accurate data on which to base judgments concerning the needs of pupils. Teachers have been justly criticized in the past because they have proceeded with their work too largely on the basis of opinion. Standardized tests have changed the situation somewhat because their use involves detailed analysis, accurate measuring, careful recording, and judgments on the basis of facts. In order to provide teachers with the necessary means of securing accurate data whenever problems arise, it is necessary that informal tests be employed more largely than in the past as a part of the regular routine of instruction. The most significant value which comes from the continuous use of standardized and informal tests is the new interest in teaching which is stimulated, and the spirit of investigation which is developed. The teacher who scrutinizes her work carefully and records accurately, becomes more openminded, more interested in adequate proofs, and less willing to follow the line of least resistance by accepting blindly all things on authority. These results cannot be secured through the occasional use of standardized tests. They develop gradually as one makes continuous use of recording devices in careful studies of classroom problems.

In the third place, informal tests are given most frequently under normal classroom conditions. This is distinctly to their advantage because measurements are of most significance when they are secured in a normal recitation. It is frequently assumed that when recording devices are used regular instruction stops. This view has its origin in the fact that standardized tests are frequently given under somewhat unusual conditions. The fact is emphasized at this point, however, that a recording device is good only to the extent that it fits into the regular work of the class and at the same time gives a series of accurate results.

Each pupil was designated by a letter of the alphabet. Entries were made on the card according to the following directions. The directions assume that the card is before the teacher turned at an angle of 90 degrees to the right from its present position.

"In scoring on the card, number the spaces from left to right to indicate the number of lines read per minute. Then for example, suppose, pupil C reads 9 lines and makes 4 mistakes. In the space numbered 9 place the letter C with a small figure 4 at the lower right-hand corner of the letter."

As the tests continued throughout the year, the teachers became more and more familiar with the methods of recording. The significance of the measurements was emphasized with each successive test. When the tabulations for the year were complete, the teachers were able to see what progress had been made by the class as a whole and by individual pupils. The requirements were simple and easily fulfilled. One of the most frequent and significant outcomes of a series of informal tests of this type is that teachers develop an interest in objective records of the results of reading instruction. It is unnecessary to add that this is one of the important results which an expert supervisor will seek to secure. It paves the way for more intensive and intelligent studies of reading, and for keener interest in the continuous use of recording devices.

Karl D. Waldo, while superintendent of the Sycamore Public Schools, organized and conducted classroom tests of reading which illustrate in detail the type of work which should be carried on regularly in each classroom. In stating the reasons for his investigation Mr. Waldo said: "As supervisor I naturally watched the reading as it developed through the phonic method and compared it with the work done under the word plan and also with the reading of the higher grades which had started their reading with no regular system. With notebook in hand, I endeavored to keep a record of the individual work of the three systems for the purpose of comparison. These records were not very satisfactory and were based on oral reading only, with no attempt to determine the comprehension by the children of the passages read, other than that which could be shown by the oral expression."¹

¹ Waldo, Karl D. "Tests in reading in Sycamore schools," *Elementary School Journal*, 15 : 251-2, January, 1915.

"When I saw how inadequate the comparison made in this way proved to be, I looked for some better method of making a record of the work. The attempts of Ayres, Courtis, Thorndike, and others to formulate standard tests suggested the formation of a plan for more scientific testing of the reading work. I also wished to test the silent rather than the oral reading, for the former is of much greater importance, not only later in school, but in future life."

Informal tests were accordingly organized by Superintendent Waldo which were simple, which closely resembled the regular work of the school, and which could be given in about an hour. The material selected for the test was in the first issue of *Current Events*, a school paper, for the year 1913-1914. The papers were passed out and the pupils were asked to proceed, when the signal to begin was given, as they would in reading any newspaper article in which they might be interested. Five minutes were allowed for the reading. When the signal to stop was given each pupil checked the last word read and wrote his name on the paper. The pupils were then asked to write a complete account of what they had read and to answer ten questions, covering the subject-matter read. The accomplishment of each pupil was then expressed in terms of his rate of reading, the percentage of the words reproduced, and a percentage grade for questions answered correctly.

The same procedure was used in all grades except that the third and fourth grades had different material and read a story of Daniel Boone in Eggleston's *Stories of Great Americans for Little Americans*. The first test was given early in the school year. After six months a second test was given in the same manner and with selections of corresponding difficulty. Check tests were given to determine the relative difficulty of the passage. The results, when tabulated, revealed three important facts² concerning the reading in Sycamore.

1. "Great variation exists in reading ability between different individuals in the same grade, and many children in lower grades are much better readers than the average of the upper grades. . . .

² *Ibid.*, p. 267.

2. "In careful reading the rate increases but slowly in school work after the fifth grade. The increase in comprehension, however, is regular from the lower to the upper grades.

3. "The third and fourth grades are places with the possibilities for wonderful development, especially in the mechanics of reading. The effects of starting the reading with a phonic system like the Ward are not evident in the reading of the upper grades. Tests tend to stimulate the work in reading, but have more effect upon the reading rate than upon the ability to comprehend what is read."

The value of such tests as an aid in supervision was summarized by Waldo as follows:³

These statements show that as a means of supervision and administration, tests like this one are of much utility. Some teachers may doubt the value of the superintendent's criticism or may feel that his opinion of the work in their rooms is unwarranted. After such a test there is no need for the superintendent to criticize. The facts found in this test were shown in the bimonthly teachers' meeting and, after discussion, were accepted by the teachers as indicative of the true situation in the reading work. Knowing the situation is the first step in the direction of effective work, and after that is taken both teachers and superintendent can, in cordial co-operation, set about improving the reading work.

The use of informal tests is not limited to elementary school classes. With increasing frequency reports are published dealing with the use of such tests in high school classes. Miss Bertha M. Smith recently reported⁴ the results of an investigation which she made in freshman classes of the high school on efficiency in assimilating reading. The passages chosen for the test related to Dickens and were taken from Long's *English Literature*. After the pupils had read intensively for five minutes they were asked to reproduce what they had read. The comprehension scores were determined in terms of the percentage of the total number of ideas which were reproduced. These records enabled Miss Smith to study the number and types of errors made by different pupils. The tests showed that those who do well in their regular school work do well in reading tests, and that those who do poorly in their daily work do poorly in the tests. The tabulated records

³ *Ibid.*, p. 253.

⁴ Smith, Bertha M. "Efficiency in assimilating reading," *School Review*, 25 : 637-45, November, 1917.

revealed the fact that both good and poor readers are ordinarily classified together and are assigned the same amount of work. In commenting on the injustice of this situation, Miss Smith said: "The teacher taking too much for granted has, no doubt, caused many a pupil to go beyond his depth. If tests were to be derived which would show what kind of mind the pupil has and, in some measure, its capabilities; then if courses were arranged in the curriculum that would take care of these variable types and would begin with them where their abilities begin, not so many pupils would complain, 'I'm not interested.' At present the good, the poor, the indifferent are grouped in the same classroom. Much of the instruction is below what the strong pupil needs, and much of it is above the weak pupil."

To remove similar difficulty the English department of the High School of the University of Chicago classifies its pupils into good, medium, and poor readers at the beginning of each year, by means of informal tests. Each group is then given assignments which are adapted to their capacities. Furthermore, the poorer sections are given training in the art of effective silent reading and study. From week to week, informal tests are given along those lines which have been selected for special attention.

Four illustrations have been given of the value of informal tests in improving instruction. Work of this type is already under way in many progressive schools. Many teachers and supervisors, however, have been somewhat slow in adopting both standardized and informal tests as essential devices for securing better results. Evidence goes to show that frequent informal tests are just as essential as books, blackboards, supplementary reading material, etc. They supply the teacher with the facts which are necessary in a scientific organization of her work from day to day.

In most cases a supervisory officer must serve as a progressive leader in stimulating this type of work. Illustrations of such studies must be presented to teachers. The value of these investigations must be emphasized. Directions for carrying on such work must be definitely outlined. Progressive teachers who are interested and who have ability to organize the work effectively must be selected for the first studies. The results of their studies should be presented to the teachers in order that they may see

more clearly the values to be derived from such studies and the methods to be employed. The supervisor must hold himself responsible for assembling scientific reports relating to the problems under consideration, and for their presentation and interpretation to the teaching staff. No opportunity should be lost to stimulate interest on the part of teachers in making detailed studies of their problems. It is because informal tests are proving of so much value in this connection that their use has been so heartily recommended in this article.

THE DEVELOPMENT OF THE UNIT PLAN OF PENMANSHIP PRACTICE

H. C. WALKER

Supervisor of Writing, St. Louis Public Schools

The waste of time and energy in teaching muscular movement writing is probably greater than the waste in teaching any other subject of the public school curriculum.

This is due to several causes, among which is the custom of teaching penmanship according to grades rather than according to the needs of the children. In arithmetic for example, the lower grades represent elementary, the middle grades intermediate, and the upper grades advanced, instruction. When applied to muscular movement writing, however, the grade in which the child is working is often immaterial. It may impose upon him subject matter which is too difficult for him to practice with correct habits; and, as a result of this, his progress in writing may be retarded rather than accelerated by his practice.

The pupil who does not know how to sit, how to hold his pen and how to use the correct movement, is a beginner in muscular movement writing, regardless of the grade in which he is working; and the only way for him to advance is to practice the elementary principles until his incorrect habits have given place to correct habits.

Too often the retardation of a child in writing is attributed to lack of interest, to absence of natural talent, or to causes other than the real one. Frequently the cause of failure is not in the child himself, but in the course of instruction which does not provide him with the kind of subject matter adapted to his particular stage of progress.

A few years after the introduction of the muscular movement into the St. Louis Public Schools, we discovered that the percentage of good writers among our eighth grade pupils was below what we desired, and that considering the time spent in practice, the results were not as good as we had reason to expect.

When muscular movement was adopted, the pupils of all grades were on the same plane as far as skill in this method of

writing was concerned: but with the continued use of the plan some of the children showed more interest in the writing than others, some showed greater skill, and to others the subject matter of the lessons was better adapted. Added to this was the steady influx into all grades of outside pupils, some of whom had not had the same training as our own pupils. The result was that, at the end of three or four years, there were in every room above the third grade three distinct groups of writers—the good, the mediocre, and the poor.

The first group was composed of pupils who wrote well in both writing practice and in daily work; the second group wrote ovals and easy capitals, but could not apply correct habits to the writing of difficult subject matter, or to daily writing; and the third could not even make ovals with correct position and movement.

We found that the needs of these three types of pupils, grouped together as a class, were as varied as would be the needs of a class of pupils in arithmetic, some of whom had advanced as far as subtraction, others as far as fractions, and still others as far as percentage.

I frequently invited the principal of the school to witness with me this adverse condition under which the teachers and pupils were working.

We found that where the teacher was adapting her instruction to the needs of the elementary pupils of her class, the intermediate and advanced pupils were wasting their time by needless review; that where she was working with the mediocre writers, the subject matter was too easy for the advanced writers and too difficult for the elementary writers; and that where she presented subject matter adapted to the needs of the advanced pupils, the mediocre pupils and the elementary pupils not only failed to derive benefit from the practice, but were actually accentuating their errors and establishing bad, rather than good habits.

Our first step toward correcting this unfavorable and discouraging condition was to organize in accordance with their needs in penmanship, three groups of writers in every room, known as the A, the B, and the C pupils. The pupils changed their seats for writing practice and each group received instruc-

tion on separate days while the other two groups studied other subjects. While this was in some respects an improvement over the former plan, it was very unsatisfactory because of loss of writing practice time for the pupils who were using the writing hour as a study period, and it was therefore abandoned.

Inasmuch as each teacher had elementary, mediocre, and advanced pupils in her room, it was decided to form three adjacent rooms into a *unit*, arrange the programs so that the writing hour of the three rooms would come at the same time, and have all the elementary pupils of the unit go to one room, the mediocre pupils to another, and the advanced pupils to another. This was done as an experiment, and the advantages of the new plan were evidenced at once by greater interest on the part of both teachers and pupils, and in the very marked improvement in the writing of the pupils of each group of the unit.

Then the promotion feature suggested itself and a definite time once a month was named for the promotion of the best writers (up to ten) of each group of the unit. The best writers of the highest group when promoted were seated in the lowest room of the unit and allowed to use the writing period for the study of other subjects, or for story reading. Some even served as demonstrators or "helpers" in the elementary group.

From this nucleus of a three-room unit, the unit plan has expanded until at present seventy of our schools have three-room or four-room units in grades from four to eight, inclusive, and other schools are adopting it as fast as the writing department can organize the work.

The three-room grouping in a twelve-room school is here used to illustrate the operation of the unit plan.

Rooms 1, 2, and 3, representing seventh and eighth grades, compose Unit 1. Rooms 4, 5, and 6, representing fifth and sixth grades, compose Unit 2. Rooms 7, 8, and 9, representing third and fourth grades compose Unit 3. The first and second grades are not included in the unit plan of practice.

Each unit is independent of the other units and the plan in all the units is the same as that here illustrated and described.

Figure 1 shows the seating at the time of organization. Class 2 is in Room 2; Class 3, in Room 3; Class 4, in Room 1. In classifying pupils Room 1 should be reserved for the poorest

writers and for those who by promotion may be advanced from Room 2 to the exemption class (Class 1). When possible it will be well to make the elementary class of Room 1 smaller than the other classes, so that the teacher may give individual aid to the pupils of this class. The reason for placing the lowest class in Room 1 at the beginning of the term is to permit this room to become the highest room at the end of the term. This condition will result from promotions. It is indicated in Figure 3.

FIGURE 1. THE SEATING OF PUPILS AT THE TIME OF ORGANIZATION.

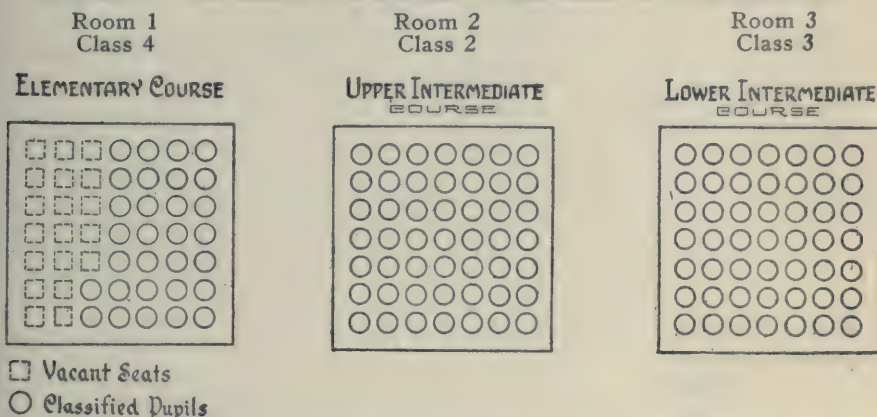


FIGURE 2. THE SEATING OF PUPILS AFTER THE FIRST PROMOTION.

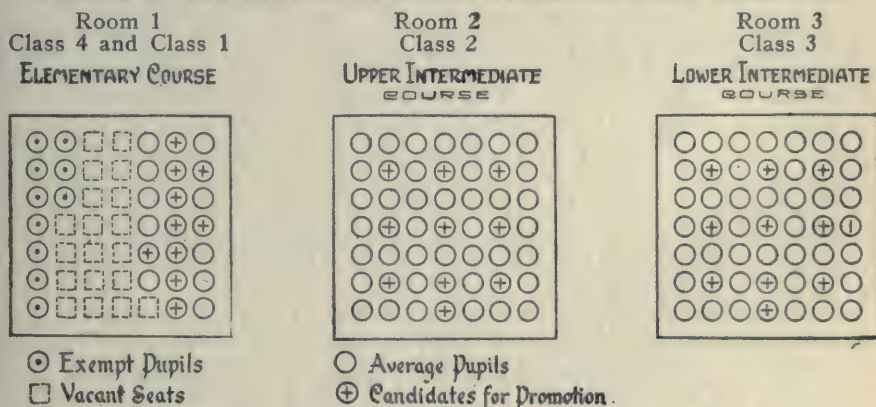


Figure 2 shows the seating after the first promotion. Ten pupils are here taken as an example of the number promoted from each room to the next higher room.

The exemption pupils, Class 1, seated on the left-hand side of Room 1, temporarily excused from penmanship practice, are permitted to use the writing time as a study period, or for supplementary reading. There is little probability that the writing of these pupils will deteriorate, because to remain in the Exemption Class, they must keep their daily writing up to a satisfactory standard.

FIGURE 3. THE SEATING OF PUPILS AFTER THEIR PROMOTION

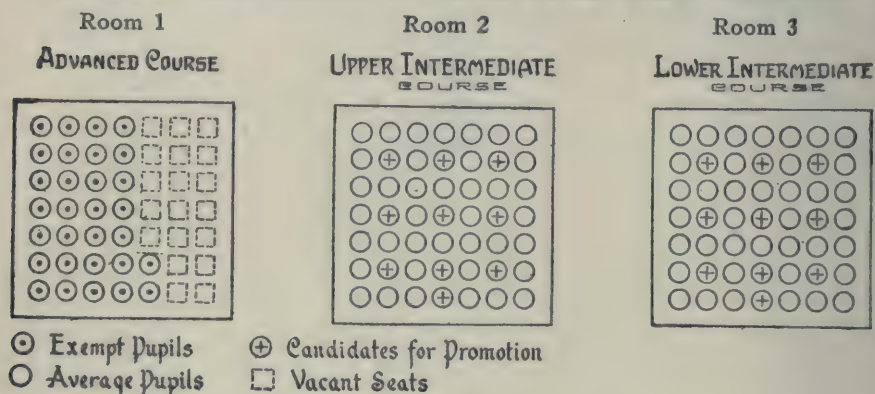


Figure 3 shows the rooms of the unit after the third promotion. The reader is to understand that a second promotion has advanced an additional ten pupils from each room to the next higher room and that now a third promotion advances a third group of ten pupils in a similar way. This makes thirty exempted pupils in Room 1 as shown in Figure 3. The elementary class (Class 4) has been eliminated and the pupils of the Exemption Class (Class 1) are ready to take up advanced penmanship practice. Now the best writers are in Room 1. Class 1 is now in Room 1; Class 2, in Room 2; Class 3, in Room 3. The rooms will retain this order until the end of the term.

The exempted pupils (Room 1, Class 1) are now ready to take up advanced practice. When the seats of Room 1 are filled, the Exemption Class is formed in Room 3. If this condition arises at all it will be found late in the school year, and the members of the Exemption Class may then be excused for the remainder of the term.

The condition existing after the exempted pupils in Room 1 have resumed practice and certain additional pupils of this room have been exempted and seated in Room 3 is shown in Figure 4.

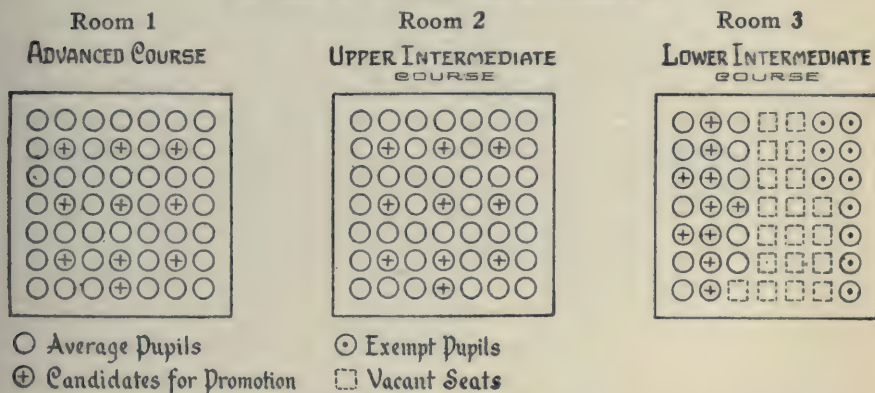
It will be seen that this plan precludes the possibility of congestion and that only those who can afford to do without the writing practice are exempted from it.

Owing to the varying sizes of desks, two-room units are suggested for buildings of eight rooms or less; three-room units for buildings of from nine to eighteen rooms; four-room units for buildings of eighteen rooms or more.

When possible the units should be formed so as to avoid making it necessary for the children to pass from one floor to another.

On a floor of five rooms, a three-room unit and a two-room unit may be organized, and on a floor of seven rooms, a four-room unit and a three-room unit may be formed.

FIGURE 4. FINAL PLAN OF SEATING



One of our principals has commented on the Unit Plan as follows:

The unit plan is advantageous because it gives the pupils who have not had sufficient elementary training in writing a chance to acquire this training, thus enabling them to strengthen their foundation, and with a stronger foundation to increase their ability to write.

The best writers, if intermingled with the poorer writers, are given so little attention that they often form incorrect habits which would have been avoided if more attention had been given them. The unit plan gives all pupils the same amount of supervision and allows them to undertake work which is on the level with their ability, thus encouraging both good and poor

pupils to put forth their utmost effort. It also stimulates interest in the poorer writers in that they work to obtain a place among the best writers, whereas formerly the poorer pupils lost this interest because they were discouraged when comparisons were made with the better writers of the class."

As a result of our experience with this new plan of practice, some interesting facts have come to light.

1. The percentage of "failures" is negligible.
2. The percentage of good writers among the boys is greatly increased and in many cases they equal or excel the girls in writing
3. The children are able to complete their formal practice much earlier in the course than formerly.
4. The supervision is greatly simplified. Formerly our writing department consisted of seven members; now it has three.
5. The teacher's problem of instruction is greatly simplified, and she is able to render not only more efficient aid to each pupil but to render it to a much larger number of pupils.
6. The penmanship is motivated without the aid of award certificates or prizes of any kind.
7. The principal's supervision of the writing is systematized. During the writing period of half an hour he may visit three rooms and find the pupils graded according to their skill in penmanship.
8. The unit plan admits of introducing into the advanced rooms of the unit composition and story writing in which the principles of correct writing may be made to function. This kind of practice furnishes the connecting link between formal writing practice and daily written work.

APPLICATION OF SCIENTIFIC METHOD IN EVALUATING THE SUBJECT MATTER OF SPELLERS*

CLIFFORD WOODY
University of Washington

Within the last decade the dominant note in educational theory has been the social point of view—that the needs of society should furnish the end points of education and should serve as a basis for the selection of subject matter. This philosophy has been urged as applicable to all phases of educational activity and the insistence on its application has resulted in many recent investigations to determine scientifically the needs of society and the consequent subject matter. In no field have these investigations been more fruitful than in spelling and the educational theorist can look with great satisfaction upon the scientifically determined lists of most commonly used words representing the social need for spelling. But the realization of his hope has merely begun, for these scientifically determined lists of most commonly used words will have little influence until they have been incorporated into the spellers which are used throughout the length and breadth of this country, since the textbook has become such a dominant force in the life of the American teacher that it tends to determine the aims, the subject matter, and the method of instruction. The use of the properly constructed textbook then is the real answer to the theorist's demand for emphasis on the social point of view.

The present study was undertaken to show the possibility of evaluating the subject matter of spellers and of scientifically selecting a book having its vocabulary made up of the most commonly used words. It represents a comparison of the Anderson list of most commonly used words—a list containing 3,105 words used 5 or more times selected from the 361,184 running words which composed the contents of 3,723 business and social letters written by adults—and the vocabularies of five widely used

*The statistical tabulations utilized in this article were made by Frank S. Salisbury and the author is indebted to him for the privilege of using them in this article.

spellers: 1) *Champion Spelling Book*,¹ 2) *Peirce Spellers*,² 3) *Merrill Speller*,³ 4) *Aldine Spellers*,⁴ 5) *The Child's Own Spelling Book*.⁵ The Anderson list⁶ was selected because it represents the latest and most comprehensive of all the attempts to select the most commonly used words and includes a very large percent of the Ayres and Pryor lists. The first three spellers mentioned were selected because they are widely used in the state of Washington; the fourth one, because the preface indicates that the vocabulary researches have been utilized to a considerable extent in the word content of the book; the fifth one, because its subject matter stands as the direct product of a study of children's writing vocabularies. In making these comparisons, the subject matter of the different spellers was the only factor taken into consideration—no attention being given to physical features of the books, provisions for review, suggestions for teachers, suggested methods, gradation of words, or price. It is true that a complete survey of the books demands that all these factors be given due consideration but in the selection of a speller it is assumed that the spelling vocabulary is of primary importance.

In making the tabulations the words of the Anderson list were transcribed onto three by five inch cards—one word being placed on a card—and the cards were arranged alphabetically, indexed, and filed. Word by word, the vocabulary of each speller was checked with the words of the Anderson list. Records were made of the words common to a particular speller and to the list, of the words that were in the list but not in the particular speller, and of those words that were in the particular speller but not in the list. Hereafter reference will be made to these groups of

¹ Hicks, W. E. *Champion spelling book for public and private schools*. New York: American Book Co., 1909. 2 pts.

² Peirce, W. N. *Peirce spellers*. New York: Ginn & Co., 1912-13. 2 bks.

³ Wilson, J. O. *Merrill speller*. New York: Charles E. Merrill, 1912. 2 bks.

⁴ Bryce, C. T. and Sherman, F. J. *Aldine spellers*. New York: Newson & Co., 1916. 4 bks.

⁵ Jones, W. F. *Child's own spelling book*. Pierre, S. D.: Capital Supply Co., 1915. 2 bks.

⁶ Anderson, W. N. *The determination of a spelling vocabulary based upon written correspondence*. Master's dissertation, University of Iowa, 1917. This list is substantially reproduced in E. J. Ashbaugh's *Iowa spelling scale for grades II to VIII* (University of Iowa Extension Bulletin, no. 34-36, September 1—October 1, 1919).

words as "overlap" words, "list only" words, and "speller only" words.

Wherever a word in the list occurred more than once in a particular book the fact was noted and in estimating the amount of overlapping such a word was given but a single count. It was originally intended to record the repetitions for the "speller only" words, but since the primary interest of the study lay in the overlapping of the spellers with the Anderson list and since the repetition of the "speller only" words would have no bearing on the problem save on total number of words in the different books, the amount of repetition was estimated as existing in the same ratio as the repetition in "overlap" of the list and a particular

TABLE I. GROSS COMPARISONS OF THE FIVE SPELLERS WITH THE ANDERSON LIST

| Spellers | Grades | Total Words in Book | "Speller Only" Words | "Over-lap" Words | "List Only" Words | Percent of Book Represented by "Overlap" | Percent of Anderson List Included |
|------------|--------|---------------------|----------------------|------------------|-------------------|--|-----------------------------------|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| Champion | 3-8 | 6,270 (5,872) | 4,846 (4,430) | 1,424 | 1,681 | 22.7 (24.2) | 45.9 |
| Peirce.... | 3-8 | 2,991 (2,862) | 2,207 (2,078) | 784 | 2,321 | 26.2 (27.0) | 25.2 |
| Merrill... | 3-8 | 5,636 (5,550) | 3,959 (3,873) | 1,677 | 1,428 | 30.0 (30.0) | 54.0 |
| Aldine... | 1-8 | 4,786 (4,436) | 2,801 (2,451) | 1,985 | 1,120 | 41.5 (44.9) | 63.9 |
| Jones.... | 2-8 | 4,532 | 2,637 | 1,895 | 1,210 | 41.8 | 61.3 |

The figures without parentheses in columns 3, 4, and 7 are only approximately true, for as was noted previously only the gross number of "speller only" words was obtained, the amount of repetition therein being disregarded. Hence a slight source of error in the figures of these columns. The figures in parentheses in the columns represent the same results on the assumption that the repetition in the "speller only" words is the same as in the "overlap" words.

speller. Record was also kept of the particular grade in which each word in each speller was taught.

The facts concerning the overlapping of the five spellers and the Anderson list are presented in Table I. The first column of this table gives the name of the speller; the second column, the grades in which it is supposed to be used; the third column, the total number of words in the book; the fourth column, the number of "speller only" words; the fifth column, the number of "overlap" words, i.e., words common to the speller and the list; the sixth column, the "list only" words, i.e., words in the Anderson list but not included in the book; the seventh column, the percent of the total number of words in the speller represented by the "overlap" words; the last column, the percent of the Anderson list incorporated in the speller.

Table I makes it evident that if the number of words included in the "overlap" (column 5), or if the percent of the Anderson list included in the "overlap" (column 8) is taken as a basis of comparison, the relative standing of the books is as follows: Aldine first, with 1,985 "overlap" words representing 63.9 percent of the Anderson list; Jones second, with 1,895 "overlap" words representing 61.3 percent of the Anderson list; Merrill third, with 1,677 "overlap" words representing 54.0 percent of the Anderson list; Champion fourth, with 1,424 "overlap" words representing 45.9 percent of the Anderson list; Peirce fifth, with 784 "overlap" words representing but 25.2 percent of the Anderson list.

To the practical minded superintendent the above comparison is fair and just for since the speller determines to so great an extent what words will be taught it is logical to insist that the subject matter of that book is best which contains the greater number or percent of the commonly used words. However, the Champion, the Peirce, and the Merrill, as shown in column 2, are intended for grades III to VIII—inclusive; the Aldine, for grades I to VIII inclusive; and the Jones, for grades II to VIII inclusive. Consequently one might expect differences in the total number of words in the spellers and more chances of overlapping in those books containing the most words. Column 3 shows that the books do vary greatly in the total number of words included—from 2,991 in the Peirce speller to 6,270 in the Cham-

pion—but comparisons with other columns in the table show little relation between the number of words in the book and the amount of overlapping. Column 7 shows that in the Aldine and Jones spellers the “overlap” is a little over 40 percent of the total spelling vocabulary while in the other spellers it ranges from 22.7 to 30 percent. Such facts tend to show the superiority of these first mentioned spellers.

To eliminate the charge that it is unfair to compare books primarily intended for grades III to VIII inclusive with those intended for more than these grades, comparisons were made disregarding the word lists in the Aldine and the Jones spellers intended for grades below the third. From Table II representing such comparison it is seen that the Jones and Aldine spellers still rank very high in the number and percent of the Anderson lists included in the “overlap” but the Merrill speller ranks higher. When comparison is made of the percent of the book total repre-

TABLE II. COMPARISON OF THE FIVE SPELLERS WITH THE
ANDERSON LIST ON THE BASIS OF THE WORD LIST FOR
GRADES III TO VIII INCLUSIVE

| Speller | Grades | Total Words in Book | “Speller Only” Words | “Over-lap” Words | “List Only” Words | Percent of Book Represented by “Over-lap” | Percent of Anderson List Included in “Over-lap” |
|------------|--------|---------------------|----------------------|------------------|-------------------|---|---|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| Champion | 3-8 | 6,270 (5,872) | 4,846 (4,430) | 1,424 | 1,681 | 22.7 (24.2) | 45.9 |
| Peirce.... | 3-8 | 2,991 (2,862) | 2,207 (2,078) | 784 | 2,321 | 26.2 (27.0) | 25.2 |
| Merrill... | 3-8 | 5,636 (5,550) | 3,959 (3,873) | 1,677 | 1,428 | 30.0 (30.0) | 54.0 |
| Aldine... | 3-8 | 3,888 (3,400) | 2,429 (1,941) | 1,459 | 1,646 | 40.0 (43.0) | 47.0 |
| Jones.... | 3-8 | 4,008 | 2,524 | 1,484 | 1,621 | 37.0 | 47.7 |

sented by the "overlap" then it is clear that the Jones and the Aldine spellers contain a much smaller percent of words of "low probability of usefulness."

Thus far all comparisons have been made with the Anderson list as a whole; but, as was pointed out by Ayres in his attempt to identify the two thousand most commonly used words, the greater the number of words the more they vary according to the subject matter under consideration, and the less reliable the conclusion about their being most commonly used. Consequently the Anderson list was divided into six parts. Part I consists of the 79 words that make up the first five deciles of the Anderson list, i.e., those words which with all their repetitions account for five-tenths of the 361,184 running words that are contained in the Anderson source material. Part II consists of the 213 words in the sixth and seventh deciles. Part III consists of the 303 words of the eighth decile. Part IV consists of the 849 words of the ninth decile; Part V consists of 1,661 words of the tenth decile which were used 5 or more times. Part VI consists of the remaining 6,118 words of the tenth decile—those words in the Anderson source material but not listed as most commonly used words.

The number of words in the different parts together with the variation of the number of times the different words were used are presented in tabular form below.

TABLE III. SHOWING NUMBER OF WORDS IN THE DIFFERENT PARTS AND THE VARIATION IN THE NUMBER OF TIMES THE VARIOUS WORDS USED

| Number of Parts | Number of Words | Range in Number of Times Used |
|-----------------|-----------------|-------------------------------|
| I | 79 | 677—11,893 |
| II | 213 | 185—675 |
| III | 303 | 79—185 |
| IV | 849 | 24—79 |
| V | 1,661 | 5—24 |
| VI | 6,118 | 1—5 |

A glance at Table III shows that each part varies in two respects, the number of words and the number of times the words

were used. It is furthermore noted that where there are few words in a part they are used many times and where there are many words they are used few times. "Few words and many times used" signifies most commonly used words, and "many words and few times used" signifies words of low probability of usefulness. Consequently that book which contains the highest number and percent in the first parts of the list will contain the most of the commonly used words.

Table IV, showing the number of words in the five spellers that are found in each of the six parts of the Anderson source material, reveals great variation among the different books in the number of words found in the respective parts. It is interesting to note that the Aldine and the Jones spellers usually contain more words in Parts I to V inclusive and fewer words in Part VI. These facts stand out more prominently in Table V, which presents the figures for Parts I to V inclusive reduced to percents.

TABLE IV. NUMBER OF WORDS IN THE FIVE SPELLERS THAT ARE FOUND IN EACH PART OF THE ANDERSON SOURCE MATERIAL

| Spellers with Book Total | Part I (79 words) | Part II (213 words) | Part III (303 words) | Part IV (849 words) | Part V (1,661 words) | Part VI (6,118 words) |
|--------------------------------|-------------------------|---------------------------|----------------------------|---------------------------|----------------------------|-----------------------------|
| Champion 5,872-6,270..... | 19 | 113 | 181 | 452 | 659 | 4,500-4,800 |
| Peirce 2,862-2,991..... | 8 | 52 | 88 | 255 | 381 | 2,000-2,200 |
| Merrill 5,550-5,636..... | 30 | 153 | 216 | 466 | 812 | 3,775-3,950 |
| Aldine 4,436-4,786..... | 79 | 118 | 206 | 618 | 946 | 2,500-2,800 |
| Jones 4,532..... | 76 | 200 | 255 | 569 | 795 | 2,636 |

Table V should be read as follows: The Champion speller, which as a book included 45.9 percent of the 3,105 words of the Anderson list, contained 24 percent of Part I, 53 percent of

TABLE V. PERCENT OF THE ANDERSON LIST AND OF THE
FIRST FIVE PARTS CONTAINED IN EACH SPELLER
(Computed from Table IV)

| Speller with Percent of Anderson List Included | Part I (79 words) | Part II (213 words) | Part III (303 words) | Part IV (849 words) | Part V (1,661 words) |
|---|----------------------|------------------------|-------------------------|------------------------|-------------------------|
| Champion 45.9..... | 24 | 53 | 60 | 53 | 40 |
| Peirce 25.2..... | 10 | 24 | 29 | 30 | 23 |
| Merrill 54.0..... | 38 | 73 | 71 | 55 | 49 |
| Aldine 63.9..... | 100 | 55 | 68 | 73 | 58 |
| Jones 61.3..... | 96 | 94 | 84 | 67 | 48 |

TABLE VI. SCORES WHEN THE PERCENTS OF OVERLAPPING GIVEN
IN TABLE V ARE WEIGHTED ACCORDING TO THE NUMBER OF
DECILES IN EACH PART

| Speller | Part I (500) | Part II (200) | Part III (100) | Part IV (100) | Part V (50) | Total (950) |
|---------------|-----------------|------------------|-------------------|------------------|----------------|----------------|
| <i>1</i> | <i>2</i> | <i>3</i> | <i>4</i> | <i>5</i> | <i>6</i> | <i>7</i> |
| Champion..... | 120 | 106 | 60 | 53 | 20 | 359 |
| Peirce..... | 50 | 48 | 29 | 30 | 12 | 169 |
| Merrill..... | 160 | 146 | 71 | 55 | 25 | 457 |
| Aldine..... | 500 | 110 | 68 | 73 | 29 | 780 |
| Jones..... | 480 | 188 | 84 | 67 | 24 | 843 |

Part II, 60 percent of Part III, 53 percent of Part IV and 40 percent of Part V, etc. This table merely serves to reemphasize the superiority of the Jones and Aldine spellers when checked against the Anderson list of commonly used words.

To give a more complete evaluation of the books on the basis of their selection of commonly used words, the parts were weighted according to the number of deciles or the part of a decile which the words therein comprise. Table VI showing the results of such weightings is derived from Table V by multiplying the percents for Part I by 5, i.e., the number of deciles in that part; the percents for Part II by 2; for Part III by 1; for Part IV by 1; for Part V by $\frac{1}{2}$. By this system of weighting the perfect scores for Parts I to V respectively are 500, 200, 100, 100, and 50; and the perfect score for the five parts taken together 950.

The total scores as given in column 7 show that the spellers should be ranked as follows: Jones, first; Aldine, second; Merrill, third; Champion, fourth; Peirce, fifth.

It should be pointed out that even though the Jones speller is given first rank according to the weighted scores, the Aldine speller has a higher score in three of the five parts. When the books were ranked according to the percent of the Anderson list included (Table I), the Aldine speller excelled the Jones by a narrow margin. When the books were ranked on the basis of the word lists for grades III to VIII inclusive (Table II) the Jones speller excelled by seven-tenths of one percent. All of these facts tend to show that little difference exists between these two spellers, and that superintendents and teachers would make little mistake in choosing either book, so far as spelling vocabulary is concerned.

On the whole, this study shows a great range in the size of the spelling vocabulary—from 2,991 words in the Peirce speller to 6,270 words in the Champion speller. It also shows the significant fact, so far as these books are concerned that little relation exists between mere size of vocabulary and the number of most commonly used words included. For example the Champion speller with 6,270 words contains but 45.9 percent of the Anderson list while the Jones speller with 4,532 words contains 61.3 percent of the list—which fact signifies that the selection of a spelling vocabulary is something more than merely collecting so many words. The study also shows that the ratio of the number of

commonly used words to total number of words in the book varies from 41.8 percent in the Jones speller to 22.7 percent in the Champion speller. In cold facts these figures mean, if we may safely argue from the basis of the Anderson list, that from 58 to 77 percent of the time devoted to spelling within the public school is spent upon words *not commonly used*. Is not one cause of the charge sometimes heard today that children are not efficient in spelling due to the fact that *we teach the children one list of words in the school room and the world expects them to spell another list when they leave the school room?*

The study shows that the Jones and Aldine spellers rank high, no matter whether considering the number or percent of the Anderson list included, the ratio of the words of the Anderson list to the total number of words in the book, or the final scores determined through the adopted system of weighting the words according to their position in the Anderson list. When only the vocabularies intended for grades III to VIII inclusive are compared the Merrill speller includes a higher number and percent of the list, but this speller is surpassed by the previously mentioned books in all other phases considered.

One must conclude then that the Jones and Aldine spellers are superior to the other three spellers under consideration as far as choice of vocabulary is concerned. This does not mean that they are superior in other respects, nor does it mean that they are superior to other spellers on the market today, especially those which have just recently come from the press. The selection of a speller, to be wholly scientific, should take account of all spellers and all aspects of the different spellers; but this was impossible in the present investigation. Such investigations require time and energy; but it can safely be predicted that until they are made we shall continue to select books, which cause children to spend their time upon words which they will probably never use, and we shall continue to select books incorporating inferior methods. Proper utilization of scientific method in the selection of spellers or other textbooks will hasten the application of the best established psychological principles and will guarantee progress in the realization of the social point of view.

THE EDUCATIONAL OBJECTIVES OF THE ILLINOIS HIGH SCHOOL CONFERENCE

H. A. HOLLISTER

High School Visitor, University of Illinois

The High School Conference which meets annually at the University of Illinois has taken as its definite task the universalizing of high school education for the state. Naturally the carrying out of such a project calls for the removal of numerous obstacles and the making of important readjustments. These preliminary steps call for much investigation along several different lines for the accomplishment of which a number of committees have been named. These committees are to operate under the direction of a central committee from the faculty of the College of Education of the university.

The studies for which committees have been chosen thus far are as follows:

1. Readjustments needed in the high school curriculum and in that of the schools leading to it.
2. Existing conditions in the state for the training of teachers and needed extension of such facilities.
3. Continuation or part-time schools—where needed, of what character, and how they are to be provided.
4. The possibilities of the all-year school—is it desirable and practicable for Illinois.
5. Provisions to be made, other than those now existing, for compulsory attendance; and the extent to which there may be necessity for subsidizing families economically incompetent in order to keep their children in school.
6. The need of extending facilities for high-school education through the organization of community high schools; the defining of an economical high school, and the problem of transportation of pupils.

At the last session of the conference a seventh committee was provided for—a committee for the study of the method of readjusting the curriculum as it affects the economics of instruction and the time and sequence elements of the problem.

It seems fitting at this time to discuss in brief detail what is involved in the way of investigation and research in these various committee assignments. Such a discussion seems all the more desirable in view of the fact that some of those interested in the movement seem not to be able to visualize the entire situation with sufficient clearness to enable them to see the necessities calling for such an extended survey. Let us take the problems, then, in the order named above.

The demand for curriculum readjustment seems to be universal, and just now there is apparently a tendency to get at the real objectives of education as a basis for such readjustment. The logical order seems to be: the determination of objectives, the selection and arrangement of materials and processes best calculated to become effective in their attainment, and the method to be pursued in the use of such materials and processes. A committee of the N. E. A. has given us a statement of objectives. But the presentation by the chairman of our conference committee, Mr. R. L. Sandwick, seems to be a better statement or classification for our purposes.

The next step, the selection and arrangement of subject matter and processes, is a far more complicated and difficult problem. We have been going on from the very beginning of our common school system with a process of unorganized accumulation of materials, both new and traditional, in a hopeless jumble. Like a retail dealer of dry goods and notions who could never manage a successful "clearance" sale we have been encumbering the shelves and attic of our educational shop with old and out-of-date stuff, much to the disadvantage of any effective "display" of the new. Is it not high time that we should take an inventory of the "goods" on hand with reference to the needs of our "trade," arranging them in the most advantageous way possible, and disposing of all non-essential and out-of-date stuff even at a sacrifice if necessary?

Such a process of weighing and testing will require much patient work; but there is no other real readjustment to be had. We shall need to apply mercilessly the tests which a clear understanding of our educational objectives should make possible. No might-have-beens or vague possibilities of the future should remain to mar our work. It is evident, then, that those who

are to participate in the testing process must first get clearly in mind the fundamental objectives of school education as determined by life needs. It will be necessary, also, that a considerable number of our best qualified teachers be available for service in this educational inventory.

The third step in curriculum reconstruction, as affecting the economics of instruction and the time and sequence relations, must, for the most part, wait upon the progress of the second step. But there is here a preliminary work to be done which may well employ the attention of such a committee until such time as it may be able to act jointly with the committee on readjustment of content of the curriculum. This preliminary study by a committee on method should result in a classification of principles of general method together with an attempt to evaluate them on a basis of accomplishment. In making such a study it will be desirable as far as possible to avoid mistaking the special application of a principle, at a given time or under a given name, for an *entirely new* principle of method.

The second study, that of teacher training, was clearly outlined in a general way in the preliminary report of the committee last November. The necessity of the situation was made very clear. There remains to be made now a careful survey for the purpose of determining the number and character of teachers necessary to carry through successfully any plan of readjustment and extension of high-school education that may result from the work of the first committee. In order to do this it is desirable, as in the case of curriculum reconstruction, to take into consideration the needs of the elementary schools as well as of the high schools. This investigation will be conducted under the direction of Dr. B. R. Buckingham, acting for the committee of which President Felmley is chairman. The study will include not only the sources of teacher supply but also their distribution, salary, tenure of service, and preparation for their work.

The third committee having in hand the problem of continuation schools is now facing a difficult problem. We already have a law on the statute books providing for the progressive and obligatory establishment of such schools until all boys and girls of ages up to and including eighteen are in attendance. But we do not know where or what kind of continuation schools

are needed; we have not enough teachers available to handle them; we do not know how the law is to be enforced; and the problem of financing such an extensive addition to our educational system has, as yet, received little or no consideration. Moreover we have not yet learned how we are to motivate the work for that considerable number of boys and girls who are out of school simply because they do not like school work as commonly organized, especially where academic subjects are concerned.

It seems evident, therefore, that there is much for this committee to do, and that it should be done as promptly as possible, if we are to save a really good piece of legislation from disastrous failure. The saner way would have been to have studied the situation first and then to have sought such legislative enactment as would fit the case.

The fourth committee has a similiar though newer phase of the problem to consider—the all-year school and its adaptability to our needs in Illinois. First of all such a committee should ascertain how this plan is working where it is already in use. It should study the various vocations in order to see how and where, as a type of continuation school, it will best serve the interests of individuals and of the state. It is possible that this type of school may be just the one needed for children of the rural districts especially.

Another problem closely allied to the third and fourth, and one which may well be handled cooperatively by these three committees, is the problem of compulsory attendance and of economically incompetent families, as assigned to committee number five. Here again methods of motivation and ways and means of effectual enforcement of law must be seriously considered. This study should include a thorough investigation of compulsory attendance laws in other states, the manner of their enforcement and the results attained. The economic problem involved in the case of families lacking adequate adult workers as bread winners needs to be carefully surveyed and the most constructive treatment possible provided.

It is useless to consider any plan for universalizing high-school education without providing schools sufficient in number as well as in character to take care of all classes of pupils who should attend. To this end the sixth committee named above was

provided. Naturally the first step should be to locate the sections where suitable high-school facilities are now lacking and to determine approximately the number and character of schools needed. This will mean a careful survey of the state for the purpose of ascertaining where there are children of high-school age—fourteen to eighteen—for whom there is at present no suitable high school available. The lack may be due either to the absence of any school near enough to be available or to the lack of any fitting type of instruction which local interests or predominant individual inclination seem to require.

The study of this committee also involves the effort to determine when a high-school organization is economically justifiable. This would imply some way of providing for a possible few who would otherwise be without high-school advantages. Thus the problem of transportation of pupils living at a distance too great for walking also falls to this committee.

It appears, then, that a very complete survey of the situation is proposed and already in progress. Purposely we have confined this to problems purely educational. The great problem of financing education is left for some one to take up as a separate study. But even such a separate study must, in some measure, be able to take into consideration many if not all of the results to come from the work we have outlined for our six committees. Certain it is that the time has come for all of us to consider our educational needs in the light of our economical limitations as related to education. We must insistently demand the elimination of waste of materials, waste of teaching talent, waste of pupil time, waste from unsanitary conditions and practices and from physical defects and disease.

Such a plan of reconstruction our present situation demands. To dally with it, to halt because of its difficulties, is only to postpone it to a later day and to make the task more difficult. It is a task whose alternative is social decadence. With all of our varied studies, therefore, there needs to be a constant effort to enlist the thought and the intelligent comprehension not only of those who teach but of all leaders of thought in all communities of the state.

One of the most gratifying elements of the above program is to be seen in the fact that any significant progress made along

the lines outlined will bring its own reward. The schools will profit by any advance made if only by one of the committees, for all these committee problems stand for real needs in our educational system. So whether we attain our objective fully or not we shall, in a measure, succeed. But if we believe, with all our hearts, in a new and better Americanism, we shall, by persisting, attain the larger goal.

Editorials

PSYCHOLOGY OF THE EDUCATIONAL "STANDPATTER"

Some are educational "standpatters" because they have never done any thinking, others because they are held captive by the thinking they once did. We are here concerned only with the latter. Sometimes he is a scholar in academic lines but has reached his saturation point. The mind has crystallized prematurely. A negativistic attitude toward the new is one of the symptoms. Such an individual becomes hypercritical toward his contemporaries. He considers their work either as cheap and shoddy or else as nothing but an amplification of something from one of the favorite authors of his youth. The pose of disdainful superiority is cultivated. He cuts himself off from his fellow workers and develops the shut-in personality.

One wonders whether such persons, notwithstanding the scholarly attainments which they sometimes possess, have not after all a fundamental intellectual defect. Perhaps it is a lack of constructive imagination that robs them of open-mindedness. Their thinking seems to have no anticipatory reference. They are unable to see how things point. They can not see the new day that is dawning in education because their intellectual heads have been set on backwards.

What a relief to turn to the educational worker who instead of specializing on the destruction of the new, just because it is new, has a never failing sense for the importance of a novel contribution, even if it is imperfect or in part erroneous. Instead of wasting his energies in the attempt to overthrow a reading test or a handwriting scale or a scale for grading English compositions he tries to develop something better. While others are proclaiming that intelligence is too complex to be measured, he devises an instrument to measure it. While others are proving the impossibility of scientific vocational selection he devises and

standardizes a trade test that works. Another in the field of school administration surveys his own school and remodels his curriculum and methods in the light of the results, while neighboring superintendents and principals are inveighing against the pretensions of measurement enthusiasts.

This is not a plea in behalf of educational science. That is on its way and will not be turned back. It is rather a plea to the young men in education, whether in the academic or the practical field, to direct their talents along profitable lines. It would also point out the reflex effects upon the worker of the lend-a-hand attitude, as contrasted with the attitude of purely destructive criticism. The one attitude favors prolonged intellectual plasticity, the other brings premature hardening of the intellectual arteries. The one means professional growth, the other professional death, not less sure because it may be insidious and slow.

L. M. T.

THE VALUE OF STANDARDIZED SILENT READING TESTS

The average silent reading rate of eighth grade pupils is approximately 240 words per minute for continuous material. Scientific investigation is revealing that this rate may be greatly increased by the application of appropriate methods. On the basis of data already secured, it is conservative to say that this rate could be increased 25 percent. As a matter of fact, it is even conservative to say that it could be increased 50 percent.

Silent reading is used as a tool in practically all subjects studied in the high school or in college. The pupil who is equipped to read at the rate of 300 words per minute has a decided advantage over the pupil who is equipped to read only 240 words per minute. It means that the pupil who reads more slowly either must spend more hours upon his work or do less work.

Of two pupils who spend an equal number of hours upon their work in high school, the one who reads at the rate of 300 words per minute will accomplish one-fourth more than the pupil who reads at the rate of 240 words per minute. This will mean that in the course of four years this pupil will do 25 percent more work, which approximates the equivalent of an extra year of schooling.

The data which we have at hand not only shows that the rate may be increased, but also that the degree of comprehension can be materially increased, perhaps to an equal extent. Therefore, not only may pupils be trained to read more rapidly but at the same time they may be trained to read with increased understanding. Thus, the pupil who has been trained to read at the rate of 300 words per minute not only will gain an extra year's schooling during his high-school course, but if his training has been of the right sort, he will actually do a better quality of work.

It appears that the use of a standardized test for measuring the rate and comprehension of pupils is one method by which the rate of silent reading and the degree of comprehension may be materially increased. At least, where such tests have been used and the results properly interpreted, these dimensions of silent reading ability have been increased. Therefore, it is not idle to say that by using standardized silent reading tests systematically, it will be possible to increase the amount of education received by pupils completing the high school by 25 percent. In fact, as we have already suggested, such a prediction is conservative in view of the facts at hand.

The fundamental resources of a state are vested in the abilities of its citizens. To develop these abilities is to enrich society. The increased wealth which may accrue to the nation from the effective use of standardized tests of silent reading may be compared with the abundance that has already resulted from improved methods in science.

W. S. M.

RESEARCH AND THE PROBLEMS OF EDUCATIONAL READJUSTMENT

The decade preceding the World War was educationally notable for the advances it brought along material lines and the external aspects of school administration. Large sums were expended for school buildings, equipment, and playgrounds. New types of schools to meet new needs were created in large number. The profession of school administration was born and brought safely through the ills and dangers which always lie in wait for an infant profession.

But the war has done three things: (1) it has left the schools financially embarrassed and has checked, for the time being, material progress; (2) it has ruthlessly exposed our educational failures; (3) it has given us a host of new educational problems.

History shows that when poverty-stricken schools face insistent demands for higher efficiency the result is likely to be a development along the lines of inner organization, method and procedure. This occurred after the Civil War; it will occur again. A new educational era is upon us. Not for fifty years have the problems of education been as fluid as they are today. It is a time for conscious self-searching and for conscious decision as to the attitude one will take toward the new problems of educational readjustment.

The period following the Civil War marked a renaissance of interest in educational method based upon philosophy and speculative psychology. Waves of Hegelianism and Herbartianism swept over the country. Rousseau, Pestalozzi, and Froebel were studied and debated. In the period which has recently begun questions of educational procedure are again coming to the fore, but with the difference that we are now seeking our solutions by the methods of science rather than by an appeal to arm-chair psychology. Everywhere tradition is being challenged by experiment. No method or program is sacred enough to escape the inquisitive and ubiquitous pedagogical test or scale. By means of intelligence scales the material with which the school works is subjected to the same cold analysis as the products of farm, factory, or mine. Nothing is taken for granted; everything must be proved. The spirit of educational research rules the day.

That much of this scientific endeavor will be ill directed or misapplied, and that much which is well directed will prove abortive because of the complexities of the problems, goes without saying. We may find that the conclusions of science no more than the conclusions of the *a priori* theorist can safely be accepted with the eyes shut. There may even arise a new type of pedant who will be entirely unable to think educationally except in terms of zero points, frequency curves, probable errors and mean square deviations!

Nevertheless educational research has come to stay. The only escape from the errors of scientific method is by a more

careful and sagacious use of that same method. The best cure for the scientific pedant will be such a development of his methods as will show finally where the boundaries lie beyond which those methods are not for the present applicable. There is little danger that opponents of the scientific movement in education will be able seriously to retard its progress. There is every likelihood that such opponents of the inevitable will lose whatever opportunities they might have had to shape the course of modern educational currents.

L. M. T.

MENTAL AND PHYSICAL AGE IN RELATION TO SCHOOL ADMINISTRATION

Compulsory education laws are driving thousands of children to the schools who, under former conditions, would not have attended. The statutes make no distinctions except those based upon chronological age and "physical or mental disability." It is assumed that when a child is six years of age, he is ready to enter the first grade. Many children are ready before their sixth birthday, many others will not be ready until later, and still others will never be ready at all. If we are going to speak in terms of age, let us use the ages which really apply. When we say that a six-year-old child is ready for school, we at once get into difficulties because of the wide range of both the physical and mental abilities of such children. We need a finer adjustment than chronological age provides. Evidently then, for school use we must have a different conception of age. If six years is the standard for entrance, we must define it as six years *physically* and six years *mentally*. Thus we have the conception (1) of physical or anatomical age and (2) of mental age. Both of these have real application to the work of the school and to the competence of children to do the work.

We can measure physical age and mental age with surprising accuracy. It is gradually becoming incumbent upon the school to do this. An x-ray photograph of the wrist bones indicates clearly their degrees of ossification. A series of standard photographs each representing average development for a chronological age will permit the matching of each child's photograph against the series so that the skeletal development of a child may be

determined in terms of *years*. This skeletal development is a reliable index of physical maturity. Almost as good an index of physical maturity, however, may be obtained from an examination of a child's teeth. Since permanent teeth should normally appear in a fixed order and at relatively fixed chronological ages, the number and character of these teeth furnish a reasonably accurate measure of physical maturity. Anyone can apply this method, for the teeth are obviously either present or absent. All one needs in interpreting the results of the count is a table indicating the ages at which the teeth appear in the child whose anatomical or physical development is normal.

Evidence shows that anatomical age corresponds much more closely to mental age than does chronological age. But although anatomical age is a much safer criterion than chronological age as a basis of school procedure, it is insufficient. In all cases it must be supplemented by a knowledge of mental age. After all, a child's success in school depends primarily upon his intelligence. The means of measuring intelligence are multiplying. The most brilliant constructive achievement of psychology and the most useful of its many gifts to education has been the Binet-Simon Intelligence Scale. This scale reveals in terms of age the mentality of children. It is a "wonderful instrument" although critics are agreed that its usefulness in measuring intelligence beyond that of the normal child of thirteen or fourteen is relatively less than it is for lower degrees of intelligence.

We are now, however, on the eve of another important application of psychology to education. The Binet-Simon Scale showed the possibilities of mental measurement in the schools. It remained, however, for the psychologists to devise a means of mental measurement which could be applied quickly, cheaply, and without the services of a specially trained examiner. Such means of measurement are now available. Some of them are being commercially distributed. The validity of these "group tests"—the extent to which they really measure intelligence—and their reliability must be more fully established. But their usefulness is now beyond question. They are not substitutes for individual tests. They give a reasonably accurate general impression. Pupils showing extreme scores should be given the Binet test.

With knowledge of the mental ages of children supplemented by knowledge of their anatomical or physical ages, we shall be better able to adjust the work of the school to that wide-ranging human variability which is everywhere present and which we cover with the term "individual differences."

We are confronted with the necessity for training large masses of pupils. Perforce we must organize them into classes. The real task of the modern school, therefore, is to reach the individual through the group. This can only be done on the basis of a grouping according to the things which really determine success in school work.

Boards of education should provide for a survey of at least the mental levels of all the children in the schools which they control. About twenty percent of the children should receive special education of some kind. The intelligence of about ten percent is so far below normal that their training must consist of activities which call for relatively little mental ability. Reading, writing, spelling, and arithmetic, although they are subjects ignorance of which is a misfortune, are also subjects which children of certain intellectual levels can never learn. If a child's mental age is only half of his chronological age, he has an intelligence quotient of 50. This will remain relatively fixed. Such a child will never become much more than eight years old mentally. He cannot learn arithmetic or even enough reading and writing to do him any good. If his intelligence quotient is 70, we may teach him some reading and writing, but it will be useless to try to teach him more in arithmetic than the fundamental operations and very simple notions of fractions. Children of intelligence quotients of 70 or 80 sometimes reach the high school. They take up Latin and algebra although they never succeed in these subjects. Indeed, we know that the high school as it is at present organized affords no place for children of such a degree of mentality.

There are, however, a vast number of children whose intelligence is below normal but not very seriously so. While these children can profit by the teaching of the common branches in their more elementary form, they should be given considerable industrial training for they will recruit the ranks of unskilled and semi-skilled labor. We need not suppose that they will be either

economically incompetent or unhappy. An increasing amount of the work of the world is being done by them; and their rewards—at least in wages—are already large, and they are increasing.

Accordingly, the task of the school is, first, to identify those whose lot it will be to serve in these capacities and, second, to give them such intellectual training as they can use and such vocational training as will fit them to perform competently the sort of work they will be called upon to do. The kinds of manual work best suited to each mental age has now been worked out in considerable detail. A woman of the mental age of seven may be taught to do well certain forms of housework, plain sewing, rug weaving, lace making, and a number of other occupations. A man of the same age may do good housework or laundry work, may do “outside work” or teaming, brush making, and bench work. In fact, the number of possible occupations for such men is rather large. Semi-skilled labor such as some employees in stores, policemen, firemen, car men, and some of the lower forms of clerical work may be and indeed should be done by persons possessing an intelligence inferior to normal, or at least not higher than normal.

At the other end of the scale of intelligence there are some eight or ten percent of the children whose gifts are so superior that they should not be subjected to the school training which is now supposed to be adjusted to the abilities of the normal child. Since, in the case of the superior child intellectual ability is predominant, he can negotiate the regular course of study in less time than can the normal child; for the regular course of study is essentially intellectual in its appeal. Many children, capable of completing the elementary school in from four to six years, are being unfairly treated and permitted to fall into loose habits of study by being required to spend eight years in the elementary school.

Not only should the gifted child progress more rapidly than the normal child, but his work should differ also in quality and method. Additional problems, more significant relationships, supplementary assignments, collateral readings—all the types of activities which we would like to carry on in the schools but for which the “minimum essentials” permit so little time—these should be prominent in the training of gifted children.

Yet special provision for gifted children is almost non-existent. Without the aid of intelligence tests the gifted child is much less frequently identified than the dull or defective child. Superior children do not call attention to themselves; they may occupy moments of enforced idleness in ways of their own choosing, but they are successful in their work. With a truly human propensity they adjust their effort to the requirements and no one may discover that they are capable of better things. This is part of the reason why we are regaled with stories of what dullards some eminent men were in their school days. If Byron and Goldsmith, Darwin and Gray were unappreciated by their teachers, we may be quite sure it was the fault of the teachers and of their means of appraisal.

The psychologist and the teacher are getting together on this question. The psychologist has simplified his methods of testing and he is training the teacher to use the simpler instruments. On the other hand, school people are realizing with conviction that they need to know and to know in terms of the new methods of measurement, the physical and mental development of the children entrusted to their care. No educational movement which goes to make up the modern trend of educational thought and practice promises more for the good of the children than does the movement which calls for the physical and mental measurement of school children and the adjustment of school procedure in accordance with the facts as revealed by these measurements.

B. R. B.

Reviews and Abstracts

TERMAN, LEWIS M. *The intelligence of school children*. Boston: Houghton Mifflin Company, 1919. 317 pp.

This book is written by an expert for the use of practical people. The number of such books is increasing. There is a cycle through which a movement passes. The idea begins by engaging the attention of a specialist to whom its implications may at first be but vaguely apparent. If he is a university man, he perhaps interests a few graduate students in his idea and writes some more or less technical articles about it—articles which practical people cannot read and which they are inclined to belittle. As the idea progresses through this stage of its development, it may begin to attract wider attention, to enlist workers in new centers, and to reveal some of its applications. Thus it comes to have a practical meaning and begins to touch the lives of a large number of people. Finally, the expert who began it may give it a broader, finer, and truer interpretation—an interpretation which he who runs may read and which cannot be ignored or undervalued by practical people.

So it is with the intelligence testing of school children. This book is written as the preface states "primarily for the grade teacher." Not so long ago the subject was engaging no attention outside of the psychological laboratory and a few institutions for the feeble-minded. The announced purpose of the book is "to illustrate the large individual differences in original endowment which exist among school children and to show the practical bearing of these differences upon the everyday problems of class-room management and school administration."

The Binet-Simon Scale, which Dr. Terman has done so much to refine and adapt for American use, is the chief means of measuring intelligence recognized in this book. Appropriately, therefore, a brief description of the scale and of the derivative measures based upon it is given in the beginning. This is followed by a general treatment of individual differences and a specific statement of their amount and meaning in the kindergarten, in the first and fifth grades, and in the first year of the high school.

It appears that a large proportion of kindergarten children could do satisfactory work in the first grade. But advancement to the first grade depends upon chronological age—upon living six years. As the author says, "the most abrupt break in the curriculum is that from the kindergarten to the first grade." Nowhere else from the first grade to the university does

advancement depend so exclusively upon age. "The kindergarten alone holds aloof, worships at the shrine of a special methodological cult, and treats its children as belonging to a different order of human beings." The need for intelligence testing in the kindergarten and of action based on the results is, therefore, especially evident.

Two first-grade classes, reported by the author, differed in average mental age by more than two years. Yet the teachers of these classes were expected to accomplish the same work. The teacher of the poorer class was in despair and her efficiency was doubted. But there was nothing wrong with the teacher. Her task was simply impossible. Differences between the mental ages of pupils in the same grade are even greater for higher than for lower grades.

As to high-school conditions, children are found to be unlikely to graduate if their mental ages are less than 0.9 of their chronological ages—in other words, if their intelligence quotients are below 90. Since "nearly a third of all children test this low or lower," the author declares that "high schools at present are in a measure 'class' schools."

Mental age is offered as a standard for grading. It is shown that if a child's mental age corresponds to the normal age for his grade, he will do average work. Incorrect grading is everywhere apparent. The dull child is promoted too much and the bright child too little. "The typical bright child, though already under age, would generally continue to do satisfactory work if promoted one or two grades."

Dullness is incurable and brightness persists. The intelligence quotient (I.Q.) is the measure of these, and it remains substantially constant. Between the first and second of 428 pairs of I.Q.'s the correlation coefficient was +0.93. Since intelligence practically reaches its maximum at the age of 16, we may by multiplying 16 by the I.Q. predict with considerable certainty what a child's ultimate mental age will be. Evidently, therefore, in the intelligence quotient we have a basis of prediction which has great value in educational and vocational guidance. It will not tell us what specific studies or occupations a child should pursue, but when the minimum intelligence required for success in each study or occupation has been learned, it will tell us what *groups* of subjects a pupil may attempt and in what *types* of work he may engage with the probability of success. Within the studies or occupations appropriate to a child's mentality choice may be made on the basis of inclination or opportunity.

Children should not be directed into occupations for which they have more intelligence than is of use. For many kinds of employment the intellectual demands are low, yet the rewards are by no means small. Present industrial developments are finding a larger and larger use for persons of inferior mentality. But in unskilled and semi-skilled occupations, intelligence above a certain level is not only unnecessary but harmful. Children, therefore, having such intelligence should be directed towards the higher types of vocations for which they are fitted. A nation's intellectual assets are priceless, and the principle of conservation will here have its most far-reaching

application—an application in which the teacher occupies the strategic position.

The author takes the attitude that unless the rank and file of teachers learn to use intelligence tests, the grading of children according to ability "will remain largely a Utopian dream." He advocates the introduction in normal schools of a half-year course of three hours per week, or of a summer session course, to prepare teachers for Binet testing. He believes, however, that teachers in service may learn the Binet procedure, and in the last chapter of his book he gives helpful directions as to how this may be done. While the supervision of a psychologist or of a person experienced in the use and interpretation of the tests is desirable, it is believed that gross errors may be avoided without such assistance. The results of the tests will never make unnecessary the securing of supplementary data from teachers and parents and the use of caution and common sense. Action, especially if based upon an assumption of feeble-mindedness or dullness, should not be taken until the results have been carefully scrutinized and compared with data as to school progress and attendance, illness, physical defects, occupation and nationality of parents, and data concerning the probable mentality of other members of the family.

Ordinarily it takes about forty minutes to test a child—somewhat less with primary children, somewhat more with the upper grade and high-school children. Where time does not permit the use of the entire scale, the author suggests abbreviations of it. He also recommends the use of some of the recently devised group tests which can be given simultaneously to an indefinitely large number of children. Group tests do not yield as dependable results as does the Binet Scale. They should be followed by a more thorough examination of children who make exceptional scores. Even the results on the Binet test are not to be taken too literally. While errors are relatively small, one must remember that the occasional error is just as serious for the child concerned as it would be if it occurred more frequently. Care in interpreting results, re-testing in the case of unusual findings, comparison with the results of educational and physical tests, with teachers' and parents' estimates, and with social conditions will prevent mistakes.

To us the most fascinating chapters of the book are those devoted to the special study of a selected number of superior children. Facts, some of which are contrary to common belief, are disclosed. The superior child is not "queer," nor physically weak. His moral qualities are likely to be superior, and his mental development is usually general rather than one-sided. His superiority is shown early in life and is permanent. Yet he is usually found in a school grade two or three years below where he belongs by virtue of his ability.

The case histories of forty-one superior children are given. They present a vivid, concrete picture of what the superior child is like. Too frequently they show how utterly these children are misunderstood and unconsciously thwarted by their teachers.

We know no more inspiring reading for teachers than these chapters on superior children. They dispel old notions and throw light upon dark places.

The conviction is forced home that such children may actually be in the reader's class, unrecognized opportunities for a type of service which every teacher worthy of the name will eagerly seek.

B. R. B.

PEARL, NORTON H. AND BROWN, H. E. *Health by stunts*. New York: Macmillan Co., 1919. 216 pp.

This little book, written by supervisors of physical education in the Detroit public schools, who also served as directors of athletic activities in the army training camps, and with the A. E. F. in France, will be welcomed, not only by supervisors and teachers of physical education, but by parents, by boys and girls, and by all who are interested in the building up of an American manhood and womanhood sound in body as well as in mind. The volume gives by verbal description and by photographic illustrations directions for performing a series of graded exercises, or better, "stunts" which are in effect a series of rough standard tests of physical ability and condition. Research men will be interested in finding many illustrations of both performance tests in which rate of accuracy are scored directly, and of ability tests where scores are determined merely by success or failure. Teachers and supervisors of physical education will value highly many suggestions in regard to methods of organizing physical education work in the schools, of conducting contests, of scoring and tabulating results, and so on, but the chief interest to workers in educational measurement is to be sought in the many possibilities the book reveals for the construction of standardized tests of physical ability. It is probable that those responsible for physical education in our schools will soon feel the pressure for proof of results which the public is exerting on every phase of public education. When the time comes for surveys and tests, this little book should prove a rich mine of test material.

For instance, in Detroit, the contestants to qualify for the Decathlon must pass a series of preliminary tests administered under standard conditions. He must be an eighth-grade boy, without organic weakness. He must be able to chin himself six times, do the "set up" sixteen times, the "floor dip" ten times, have a grip of at least sixty pounds and a lung capacity at least two and six-tenths times his height in inches. Why should there not be such age-grade standards of physical fitness all along the line? When the product of all physical training is checked against suitable objective standards it will become much more effective than it has in the past, and the physical fitness of the average American will be assured.

S. A. C.

BOBBITT, FRANK M. *The curriculum*. Boston: Houghton Mifflin Co., 1918. 295 pp.

During the last ten years increasing attention has been paid to the construction of curriculums in school subjects, particularly in the fields of spelling and language; but no systematic treatise has been published during that time until the appearance of Professor Bobbitt's book.

The book is refreshing in its beauty and simplicity of style. It reads easily and speaks directly to the point in dealing with the topics under discussion. These are grouped in six sections: ends and processes, training for occupational efficiency, education for citizenship, education for physical efficiency, education for leisure occupations, and education for social intercommunications.

In the first section the groundwork for treatment of the remaining groups is laid. In this section a valuable distinction is made between experience gained by undirected study on the play level and that obtained by directed study on the work level. Of particular significance is the author's claim that, after experience has been gained without direction by children, the curriculum that follows this "*aims at those objectives that are not sufficiently attained as a result of the general undirected experience.*" Or, to paraphrase, after a child has picked up what he will by himself, the school should discover his shortcomings and correct those. Thus, for instance, the author says, in language, the school curriculum should be based upon the language errors of children. In vocational education it should be based upon occupational deficiencies.

It is not entirely clear in my mind whether undirected experience is a part of the curriculum or not. The author does not discuss the point; but this is probably an oversight. Obviously, it is not practicable to let children pick up experience everywhere. If a child is turned into a library to browse, it is necessary that the books be selected upon some basis. If he is allowed to run at large, the boundaries must be set in some way: the sawdust of a saloon floor ought certainly to be taboo for a growing boy. And, if this "undirected experience" is really selected and, therefore, part of the curriculum, errors and shortcomings will hardly constitute an adequate basis for selection. An interesting argument, at least, could be made on the subject.

It is unfortunate that the author does not show the applications of the shortcomings concept to the construction of directed curriculums in citizenship, physical education, literature, and the foreign languages. But this is probably not an omission because the text is a "first book" and deals largely with the functions of the activities mentioned, rather than with the content of the curriculum necessary for the realization of each.

The book is timely and will be of fundamental value to any student of the curriculum.

W. W. C.

PUBLICATIONS RECEIVED

- Baker, Franklin T. and Thorndike, Ashley H. *Everyday classics*. New York: Macmillan Co., 1918-1919. 6v. (Third-Eighth Readers.)
- Baker, Franklin T. and Thorndike, Ashley H. *The teaching of reading: a manual to accompany everyday classics, third and fourth readers*. New York: Macmillan Co., 1917. 93 pp.
- Beard, Charles A. and Bagley, William C. *The history of the American people*. New York: Macmillan Co., 1919. 686 pp.

Bibliographies

A BIBLIOGRAPHY OF STANDARDIZED TESTS FOR THE HIGH SCHOOL¹

WALTER S. MONROE,

Assistant Director, Bureau of Educational Research, University of Illinois

In addition to standardized tests designed to measure the abilities of pupils in high-school subjects, certain tests designed for use in the elementary school have been given to high-school pupils. In assembling this bibliography an elementary school test was included if a published report of its use in high school was found. Doubtless other elementary school tests have been or could be used with profit in the high school.

A special effort was made to have the list of tests designed for use in the high school include all of which an account had been published. A few tests have been included which are in the process of construction. Although it may be that some tests escaped notice, this list is more complete than any which has been previously published. In a few instances it was difficult to determine whether the methods used in constructing a list of questions or exercises were sufficiently scientific to justify its inclusion in this bibliography. In general, doubtful tests were included. For this reason the reader is cautioned against attaching undue importance to the fact that a test appears in this list. Before a test is selected for use, careful consideration should be given to both the method of its derivation and the results of its use. When accounts of these are lacking, the advice of competent authorities should be sought.

The list of references for a test includes the account of its derivation, if it has been published, and the more important accounts of its use. The brief annotation indicates the character of the reference.

The publisher from which a test can be obtained is given when it could be ascertained. The Bureau of Educational Research of the University of Illinois carries in stock a number of tests which are published elsewhere, and for most of these the Bureau has devised class record sheets or other accessories. For this reason the Bureau of Educational Research, as well as the original publisher, is given as a source from which such tests may be obtained. A few of the tests listed are not available for distribution. In these cases the address of the author is given if it is known. All prices are omitted because the instability of the printing industry prevents announcing them permanently.

Under the head of "General References" there is given a selected bibliography which deals with the general topic of educational measurement. Some of the references

¹ This bibliography was prepared with the assistance of Miss Margaret Doherty, Library Assistant in the Bureau of Educational Research, University of Illinois.

pertain to tests for use in the elementary school, but they will also be found helpful by those interested in the high school field.

A section is devoted to mental tests. This particular field is developing so rapidly at this time that it is impossible to give a complete list of the tests that have been devised or that are in the process of derivation. For this reason the reader must regard this particular list as tentative and incomplete.

GENERAL REFERENCES

Chapman, James C. and Rush, Grace P. *The scientific measurement of classroom products*. New York: Silver, Burdett and Co., 1917.

For the most part this book deals with tests for use in the elementary school. There are three general chapters which will be of interest to high school teachers.

Courtis, Stuart A. *The Gary public schools: measurement of classroom products*. New York: General Education Board, 1919.

Besides being a report of the use of a number of standardized tests in the Gary public schools, including the high school, this volume is an important contribution to the theory of educational measurements. The validity and reliability of the tests used in the survey are treated.

The measurement of educational products. (The Seventeenth Yearbook of the National Society for the Study of Education, Part II). Bloomington, Illinois: Public School Publishing Company, 1918.

This volume of the *Yearbook* is a compilation by members of the National Association of Directors of Educational Research. A complete bibliography is given.

Monroe, Walter S. "Improvement of instruction through the use of educational tests," *Journal of Educational Research*, 1:96-102, February, 1920.

This article presents a procedure for interpreting class scores and distributions with special reference to the Courtis Standard Research Tests, Series B.

Monroe, Walter S. and De Voss, J. C. and Kelly, F. J. *Educational tests and measurements*. Boston: Houghton Mifflin Co., 1917.

This is a general treatise upon educational tests and their use. Chapter VII is devoted to high-school tests.

Proctor, W. M. "Psychological tests employed in the vocational guidance of high school pupils." (To appear in the first volume of the *Journal of Educational Research*.)

Starch, Daniel. *Educational measurements*. New York: Macmillan, 1916.

This book contains two chapters upon school marks. Most of the tests which were in existence at the time the book was written are reproduced. A number of them are designed for high school subjects.

Terman, Lewis M. *The intelligence of school children*. Boston: Houghton Mifflin Co., 1919.

This book emphasizes the need for the use of intelligence tests in the classification and guidance of pupils.

Terman, Lewis M. *Measurement of intelligence: an explanation of and a complete guide for the use of the Stanford revision and extension of the Binet-Simon intelligence scale*. Boston: Houghton Mifflin Co., 1916.

I. COMMERCIAL SUBJECTS

National Business Ability Tests

1. Test on tabulation (mental alertness)
2. Reproducing instructions
3. Test on invoicing
4. Tests on fundamental arithmetic:
 - a) Fractions
 - b) Percentage
5. Test in business arithmetic
6. Tests in English
 - a) Spelling, elementary and advanced
 - b) Grammar and punctuation, elementary and advanced
7. Elementary test in letter writing
8. Test on answering letters (advanced)
9. Stenographic tests (transcription and typewriter copying)
10. Test on copying for the mimeograph
11. Test on addressing envelopes with a pen and on filing
12. Ayres' Handwriting Scale used to judge handwriting

The nature of the above tests is indicated in a general way by the titles. For a detailed description consult Chapter v of the reference.

For each of the tests there are "two series" or two forms. It is intended that the first series be used when the test is given for the first time and the second series is to be used when it is desired to repeat the test. The two series are intended to be equivalent in difficulty.

Publisher: National Associated Schools of Scientific Business, Inc., Chicago.

Reference: Cody, Sherwin. *Commercial tests and how to use them*. Yonkers-on-Hudson, New York: World Book Company, 1919.

Part One contains the following chapters: Schools and business employment, National Business Ability Tests, principle of scientific tests, national tests in the classroom.

Part Two consists of the following chapters: Two series of National Business Ability Tests (tests are reproduced, with directions for giving and scoring), measuring classes and teachers. Appendix is devoted to the report of use of the tests in classroom.

Rogers' Stenographic Tests. This is the combination of tests which Mr. Rogers found to be useful for identifying those persons who possessed the qualities necessary for a successful stenographer:

1. For dictation
 - a) Hard Directions
 - b) Opposites
 - c) Form Substitution
2. For grammar
 - a) Hard Directions
 - b) Opposites
 - c) Mixed Relations
3. For typewriting
 - a) Verb Object
 - b) Color Making

Reference: Jaques, M. P. "Mental tests for typists and stenographers," *Industrial Management*, 58:145-46, August, 1919.

A discussion of an experiment with the Rogers' Stenographic Tests in the Brooklyn branch of the Charles Williams Company.

II. DOMESTIC SCIENCE

Bowman and Trilling's Informational and Reasoning Tests in Textiles and Clothing.

"These tests have been constructed by Leona F. Bowman and Mabel Trilling of the University of Chicago and will be described in a forthcoming number of *Supplementary Educational Monographs* published by the School of Education, University of Chicago."

Murdock's Scale for Measuring Certain Elements in Hand Sewing.

Publisher: Bureau of Publications, Teachers' College, Columbia University, New York City.

Williams and Knapp's Scale for Measuring Skill in Machine Sewing. The method of deriving this scale is similar to that used in Thorndike's Handwriting Scale.

III. DRAWING

Manuel's Suggested Series of Tests for Studying Talent in Drawing. On the basis of experimental testing of nineteen persons who had marked ability in drawing, six of whom were high-school students, Dr. Manuel suggests the following set of diagnostic tests:

1. Tests of the elementary ability to represent by lines and areas, figures and objects directly observed. Drawing a horse (memory) and a toy wagon (object).
2. Tests of general intelligence (Stanford-Binet revision) with special reference to:
 - a) Memory for Digits
 - b) Designs
 - c) Sixty Words
 - d) Clock Test
 - e) Vocabulary
 - f) Fables
 - g) Code
 - h) Inclosed
 - i) Paper Cutting
 - j) Logical Memory
3. Tests of ability to discriminate differences in visual magnitudes (Manuel's Test in Discrimination of Visual Proportions, Manuel's Series).
4. Tests of visual acuity and normality of color vision.
5. Tests of ability to observe visually:
 - a) Whipple Cancellation Tests—4 forms
 - b) Rossolimo Observation Test
 - c) Whipple Description of a Stamp
 - d) McDougall Spot Pattern
6. Tests of memory for visual forms especially deferred memory:
 - a) Rossolimo Recognition of Lineal Figures

- b) Rossolimo Recognition of Colored Figures
- c) Rossolimo Recognition of Pictures
- 7. Tests of ability to manipulate spatial forms.
 - a) Thurstone's Hand Test
 - b) Thurstone's Spatial Relations Test
 - c) Punched Holes
 - d) Rugg's Painted Cube Test
- 8. Tests of esthetic judgment:
 - a) Thorndike's Tests of Esthetic Appreciation
 - b) University of Illinois—Esthetic Judgment Test.

Reference: Manuel, H. T. *Talent in drawing: an experimental study of the use of tests to discover special ability.* (School and Home Education Monographs No. 3). Bloomington, Illinois: Public School Publishing Company, 1919.

This monograph is devoted to a description of the tests and an analysis of the results obtained.

Rugg's Scale for Measuring Freehand Lettering for Use in Secondary Schools and Colleges. It consists of a series of eight samples of freehand lettering, arranged in the order of increasing merit. It may be used in measuring the efficiency of a student's work in freehand lettering.

Reference: Rugg, H. O. "A scale for measuring freehand lettering," *Journal of Educational Psychology*, 6:25-42, January, 1915.

Derivation, function, and experience with the scale is discussed.

(To be continued in the March Number.)

News Items and Communications

The George Peabody Foundation has recently sponsored a movement that is full of promise for the future cultural and educational relations of America with foreign nations. This is no less than the effective organization of all correspondence between pupils studying modern foreign languages in American schools and those similarly engaged in the schools of France, of Spain and South America, and of such other lands as shall eventually be included. Hitherto such letter exchange, while of great value in individual instances, has been ephemeral, spasmodic, and disorganized. It has not realized its full potentialities. The Peabody Foundation has set aside one of its buildings on the campus of the Peabody College for Teachers at Nashville as the national American office and has appointed Dr. Alfred I. Roehm as national director. With the cooperation of state and community chairmen all over the country the work of international educational correspondence is rapidly being organized.

The first branch established was the French-American Bureau. In France Monsieur Charles Garnier of the French Ministry of Education has

**International
Educational
Correspondence**

charge of the central French office in Paris. Lists of pupils in French and Belgian schools, who have been selected as especially fitted to inaugurate the correspondence, are sent to the American office at Nashville where the details of exchange with American schools are arranged. A similar Bureau to cover correspondence with Spanish-speaking lands has been authorized by the United States Department of State, and this will be organized immediately. Other bureaus to include other foreign languages will follow in due course.

The plans of the Peabody Foundation have received worldwide approval. The possibilities of liberalizing influences and contacts are without limit. To quote from Mr. Claxton's article in *School Life* for September 15, 1919: "Linguistic training will not be the only educational end served. Along with the letters there will be a fine exchange of historical, artistic, geographical, manufactural, commercial, and home-life material and information clippings, picture postals, kodak views, etc., leading up to the deepest exchanges of human sympathies and ideals that will reinforce international, good-will." In addition to the above article of the Commissioner of Education the reader is referred for details to a description of the purposes and plans of the Foundation in *The Modern Language Journal* for November, 1919, pages 73-76 by the present writer. Another shorter article of interest may be found in *School and Society* for September 27, 1919, page 377. It is manifest that pupils cannot at once write letters of very great value in the foreign language. Therefore it is suggested that the initial letters be in the pupil's own tongue. Thus each school will gradually have a number of model letters in the foreign idiom. Careful study of these models will be far more stimulating than the usual arid exercise of the average composition manual. The pleasure of actuality is here of great value. This has been the universal experience of teachers who have tried letter exchange in lieu of set composition. In due course the pupil will substitute the foreign tongue for his own. Each letter received is a challenge to send in return an equally good, or a better, answer. It is of course deemed wise that boys shall correspond with boys, and girls with girls, since mutual interests are thereby better utilized. While the burden upon the teacher will doubtless be somewhat greater at first than under the usual conditions of composition work, yet the same stimulus exists for teacher as well as pupil. It is even quite conceivable that the net result may prove a lessening of the teacher's task rather than an increase, since pupils invariably do better that which they like to do. At present no composition manual based upon epistolary style exists, but such books will soon appear in response to an evident demand, and they can greatly help in this important work.

It is the aim of the Peabody Foundation to establish this enterprise upon a self-supporting basis as soon as possible. Therefore communities and schools are asked to contribute to a permanent endowment fund in one of two ways. By the first or "quota method" each community will pay two dollars for each one thousand inhabitants. This single payment will entitle all persons in the community to all services of the Foundation forever. By the second or "fee method" every pupil enrolled for international correspon-

dence will pay an annual fee of ten cents. Already there have been enrollments under both these methods from every state in the Union, and the interest is rapidly spreading. Until this permanent endowment has been secured, the Peabody Foundation will bear the expense of the enterprise. Later it will be the trustee of the endowment and will render annual reports to the United States Bureau of Education.

The liberalizing effects of such a system of international correspondence are certainly beyond all doubt. All friends of education see in this plan very great possibilities of good and will be eager to contribute to its success.

THOMAS E. OLIVER
State Chairman for Illinois.

**Manual Arts
Conference,
University of
Cincinnati**

The tenth annual conference of specialists in industrial education engaged in training teachers in industrial and manual arts in institutions located in the Mississippi Valley was held at the University of Cincinnati, December fourth, fifth and sixth, 1919. About forty-five men representing approximately as many institutions were in attendance.

These conferences are called from year to year by the United States Commissioner of Education at the request of the group, there being no permanent organization. For the past ten years Dr. William T. Bawden, Assistant to the Commissioner, has acted for the Commissioner in the matter arranging time, place, and program.

The following papers were presented at the Cincinnati meeting. The program was planned to allow generous time for discussion by the group. In fact the meeting is looked upon primarily as an opportunity for the exchange of views by those present.

"An experiment in developing a course in foreman training" by Charles R. Allen, Agent for Industrial Education, Federal Board for Vocational Education, Washington, D. C.

"Progress in the development of plans for preparing teachers of industrial subjects" by Professor James McKinney, Director of the Chicago Center of Teacher Training Department of Industrial Education.

"The itinerant teacher of subjects in village and rural schools" by Professor C. S. Van Deusen, State Normal School, Kent, Ohio.

"Provision for the training of vocational teachers in the army" by Professor R. W. Selvidge, Educational Adviser, War Plans Division, War Department, Washington, D. C.

"Opportunities for service in the field of occupational therapy" by Lieutenant Colonel Mattock.

"Training teachers of vocational and industrial work not federally aided" by Professor C. A. Bowman, Stout Institute, and W. E. Roberts, Cleveland, Ohio.

The program of the meeting also included a report of the committee on practice teaching which was concerned with plans. One of the best of these was submitted by Professor A. F. Siepert, Bradley Institute, Peoria.

A Chance to Test the Rural Schools At the annual meeting of the Association of County Superintendents of the State of Illinois held at Springfield on December 29 a committee of five was appointed to facilitate cooperation with the Bureau of Educational Research at the University of Illinois. This committee is composed of the following members:

G. O. Smith (Chairman), Bureau County, Princeton

Mrs. Lois Coombes, Moultrie County, Sullivan

F. E. Crawford, Fayette County, Vandalia

Charles H. Watts, Champaign County, Urbana

J. W. McKinney, Williamson County, Marion.

The immediate objective of this committee is to increase the use of standardized tests in the rural schools. Already some superintendents have given one or more standardized tests in all the schools of their county but it is the feeling of the county superintendents as expressed through their state association that the use of standardized tests should become general throughout the state.

* * * * *

The difficulty of reaching individual rural schools from one center such as the state Bureau of Educational Research is at once apparent. However, the relative isolation of the rural teacher creates a situation in which a standardized test can be even more helpful than to a teacher in a city school system who is working under close supervision. It will give her definite objectives for the pupils of each grade and, since in many cases she has a small number of pupils, it will be possible for her to make a relatively complete diagnosis of the shortcomings of her pupils. If these valuable instruments for measuring instruction are to become generally used in rural schools, it is incumbent upon county superintendents to be active in bringing them to the attention of rural teachers.

The action of the county superintendents of Illinois represents a progressive movement. In many states particular county superintendents have been using standardized tests in their schools for some years, but this is thought to be the first instance of concerted action by all of the county superintendents of a state.

The Federal Bureau and Educational Research Commissioner P. P. Claxton of the United States Bureau of Education called a meeting of representatives from a number of the larger state universities at the Hotel Statler, St. Louis, Missouri, January 2 and 3, 1920. This meeting was attended by the following representatives:

University of California: R. S. French

University of Illinois: B. R. Buckingham and Walter S. Monroe

University of Kansas: F. J. Kelly

University of Michigan: Guy M. Whipple

University of North Carolina: L. A. Williams

University of Pennsylvania: J. H. Minnick

University of Texas: C. T. Gray

In addition to Commissioner Claxton his assistant Wm. T. Bawden was in attendance.

Recently, upon the invitation of the Commissioner of Education, research stations have been established at a number of universities. In addition to the institutions represented at the meeting, the following appointments have been made: University of Iowa, University of Minnesota, University of Virginia, and University of Washington.

In a circular issued September 15, the purpose of the plan of establishing research stations was stated as follows:

(1) To promote scientific research in education, and to make the results of such research available to students of education.

(2) To utilize in a national way such special resources and facilities as may be placed at the disposal of the Bureau of Education by the cooperating institutions.

(3) To coordinate and stimulate the educational forces available in the cooperating institutions.

(4) To bring the Bureau of Education into closer relations with the educational agencies of the country, and thereby to promote the Bureau's understanding of educational needs and problems, and to extend its range of effective service.

This move upon the part of the Bureau of Education is significant. It means first a closer identification of the Bureau with the schools of education in the larger state universities. It also means that the United States Bureau of Education is identifying itself with the rapidly developing recognition of and provisions for educational research. Within a period of little more than five years there has developed a new type of department in the field of education, namely a department which has for its particular function the stimulating and directing of educational research. The larger portion of this research is carried on in cooperation with public schools. The recognition of this development by the United States Bureau of Education is particularly significant at this time when an effort is being made to secure a larger recognition for education throughout the country.

From the standpoint of the Bureau of Education, this move will mean that it will assume a larger place in the educational activities of the country. It is a lamentable fact, but nevertheless true, that the recognition given to education by the government is far less than that accorded to this field by most governments. If the establishment of research stations of the Bureau does no more than increase the prestige of the United States Bureau of Education, the project will be justified.

National Association of Directors of Educational Research

(E. J. ASHBAUGH, *Secretary and Editor*)

All aboard for Cleveland!

There is every reason why this meeting should be the banner meeting of the National Association of Directors of Educational Research. The program has been prepared earlier than ever before. A larger amount of time will be given to the discussion of the problems of research men. Every member is urged to go through the program carefully and then decide that he cannot afford to miss this meeting.

Special attention is called to the closed meeting on Monday, February 23. This is to be a meeting for members only and will be devoted to intimate discussion of problems of research. Each member should come to the meeting prepared to submit a small list of carefully defined problems for this discussion. From the standpoint of the membership it should be the most valuable day of the whole week.

In order to get off right foot first with the Journal, your secretary has sent out a notice to the members asking their intention concerning continuing membership this next year. Almost 100 percent replies have been received and with but a single exception the statement has been affirmative. The one negative vote is from a former member who is not now engaged in research work. A few have pledged themselves for membership for life. This is the spirit that wins.

Since the writing of the notes for the last issue, word has reached us that the five million dollar bond issues that were then pending in Omaha and Oakland were both carried by a good substantial majority. There is every evidence that the great American public is behind the public schools more emphatically than ever. It certainly vindicates also the impression which many of us have had that people are perfectly willing to tax themselves for the education of their children when the facts showing conditions and needs are placed before them in a clear, concise manner. Mr. Smith and Mr. Dickson should both receive the unanimous commendation of the members of the Association for the success of the work.

Mr. Shapleigh, Research Secretary, Buffalo Public Educational Association, Buffalo, New York, should be congratulated upon the issuance of an educational journal for the city—*School and Community*—two

issues of which have reached the secretary's desk. Both represent a very high degree of excellence in both content and composition. Other school communities of large size can certainly afford to use similar means of keeping the public well informed. Wise publicity would very quickly drive out of existence the contemptible type of cartoon of schools and school conditions which so often disgraces the pages of our public press.

We are glad to welcome to membership in our Association: Charles L. Harlan, Department of Educational Measurement and Standards, Lewiston State Normal School, Lewiston, Idaho, and John K. Norton, Assistant Director, Bureau of Research and Guidance, Oakland Public Schools, Oakland, California.

If a member knows of any directors of educational research who are not now members, he will confer a favor upon the Association by recommending to the secretary such of them as he believes to be acceptable for membership. If this notice should come to the attention of any one who is not a member of the Association but who is engaged primarily in educational research, a letter to the secretary will bring him information concerning eligibility to membership and the general scope and benefits of the Association.

Please remember that no further issues of the news bulletin will be made. New items from the various members are to be given publicity through this department. The secretary urges the members to feel under personal obligation to send him items concerning their work. Whenever you have completed a bit of research which you feel has been helpful to yourself or to your community; whenever you start something which is out of the regular routine; whenever you have a piece of projected work to which you think other members might contribute some comparable data, please spend a few minutes time and a postage stamp in getting these data to the secretary.

At the closed meeting of the Association which will be held on Monday, February 23, topics of especial interest to the members will be taken up. None but members will be expected to attend unless they are particularly invited to do so. Among others, Dr. Thorndike will address the closed meeting by invitation. He is going to talk on "Measurements of instruments of instruction" (i.e., textbooks). He intimates that he will "start something."

A number of topics of general interest to members will be proposed and leaders of discussion will be appointed. The following topics have been suggested: the improvement of instruction by the use of standardized tests, the evaluation of tests, group intelligence tests. If any of the members wishes to have a topic discussed he should communicate the fact to the president at once. While the meeting will not be of a formal character it will nevertheless be definitely planned; and it may be difficult to secure the discussion of a topic unless a place has been arranged for it.

For the use of the closed meeting the management of the Hotel Statler has placed at the disposal of the association a committee room having a

seating capacity of forty. Adjoining this room there will be a private dining room where luncheon and dinner will be served to the members. The use of the committee room is given to the association gratis; but the members will be expected to eat luncheon and dinner in the dining room assigned to the Association. They would anyway!

At this meeting a preliminary report of the committee on standards will be rendered. The report will be open for discussion. The closed meeting will also afford an opportunity for an informal discussion of business matters. This will enable us to dispose of them quickly at the business meeting Thursday evening. If any member has an item of business which he desires to have debated he may bring it up at the closed meeting.

The following papers are announced for the public meeting to be held Thursday morning and afternoon, February 26:

W. W. Theisen. "Provision for individual differences in the teaching of reading."

H. A. Brown. "Formulation of method in reading in the light of recent investigations."

William S. Gray. "Significant tendencies in the diagnosis of reading accomplishment."

Edward L. Thorndike. "Intelligence tests for college entrants."

F. J. Kelly. "Results of three methods of teaching the fundamentals of arithmetic, measured by the Courtis Tests."

Cecile White Flemming. "Reading ability and intelligence."

S. A. Courtis. "Report of committee on standards."

Leonard P. Ayres. "The comparative effectiveness of state school systems."

Arthur W. Kallom. "Reproduction as a measure of reading ability."

J. W. Richardson. "The campaign method in elementary education."

As will be noticed, the program is largely grouped around the ideas of reading and intelligence testing.

The annual dinner will take place on Thursday evening at 6:30 o'clock at the Hotel Statler. Price per plate \$3.00.

Journal of Educational Research

Editorial Office, Urbana, Illinois

Editor-in-chief B. R. Buckingham, University of Illinois

VOLUME I

MARCH, 1920

NUMBER 3

A PROPOSED INDEX OF EFFICIENCY IN TEACHING UNITED STATES HISTORY

B. R. BUCKINGHAM

University of Illinois

When a supervisor enters a classroom to inspect the work and rate of the teacher, he carries with him a set of more or less variable personal standards. He may also fortify himself with a schedule of qualities on the basis of which he will perhaps believe that he is scientifically evaluating the work of the teacher. We have a number of score cards which set up such captions as personality, voice, neatness, method, control, and the like. These are by no means satisfactory. Teachers possessing these qualities to a high degree may nevertheless be poor teachers. Personality as a producing agency, neatness functioning, voice used to get results, method which does the business, and control which facilitates work—these qualities, in other words, not as possessions of the teacher but as productive of changes in pupils—are touchstones of good teaching.

We must, therefore, seek for teaching efficiency in the changes produced in children rather than in the qualities possessed by teachers. We are not, however, to suppose that we have greatly simplified our problem when we have thus defined it. The results of teaching are numerous and far reaching. We can never hope to measure all of them, for they multiply to the end of life.

Yet we are not without means of measuring with a considerable degree of accuracy the efficiency of teaching. What we need to do is to find some one major result the simplicity of which makes its apprehension relatively easy and the relationship of which to

the totality of result is close. If there were a perfectly definable and generally recognized relation between the length of the left forefinger and the stature of a human being, and if it were difficult to measure stature but easy to measure finger length, then the height of human beings might very probably be expressed not directly but indirectly through the measurement of the left forefinger. Countless uses are made of the principle of directly measuring a magnitude in order that another magnitude related to it may be indirectly measured. A printer often counts the material in a pile by measuring its height or weight. Taxes have been assessed according to the number of windows in the houses because the value of the house and hence the wealth of its owner were thought to be roughly proportional to the number of windows. We weigh the air by measuring the height of a column of mercury which it will sustain; and the weight of the air is itself functionally related to the weather conditions which we are thus, in turn, able to estimate. Again, we directly measure the expansion of a fluid in a tube and thereby indirectly measure temperature.

In the matter of teaching it is necessary in the first place to seek our index of efficiency among the results of the teaching. In the second place, it is desirable that this index shall be of a simple and easily measurable nature.

Some of the results of teaching are being measured objectively by means of standardized tests and scales. Opposition to educational measurement as such is all but silenced. Not many voices are now being raised against it on principle. Numerous critics, however, assail its practical applications and they do so at many points and in many ways.

It is not unusual, for example, for the critic to contend that certain subjects do not lend themselves to measurement. This assumption is a dangerous one, and when the critic is so rash as to make his objections apply to a particular subject he is likely to be refuted by the next educational periodical or bulletin. In 1913, when measurement had been applied for two or three years to handwriting and arithmetic, it was pretty generally believed that these two subjects might be susceptible to such treatment, but that one could go no further. In particular, it was roundly asserted by a speaker, and cordially approved by his hearers, that spelling was a subject that would never yield to measure-

ment. Yet everyone knows that to measure the spelling ability of pupils is one of the first steps of the novice in measurement. Similarly, nobody thought that ability in reading could be measured, yet the sale of millions of copies of reading tests now assures us that its measurement is at least being widely attempted, and that people believe they are securing reasonably reliable results.

In the case of United States history, there is evidently a phase of the subject which lends itself admirably to measurement. The method of expressing the difficulty of the elements in any test has been carefully worked out. It has, therefore, become perfectly feasible to apply this method to any subject, the difficulty of whose elements can be determined. One definition of difficulty—not the only one, to be sure, but one which is especially significant—is the proportion of correct responses to the task in question. When did Columbus discover America? Who was president of the Southern Confederacy? What two probable European ventures was the Monroe Doctrine set forth to forestall? These are matters whose difficulty on the basis of the proportion of correct responses may be as accurately determined as the difficulty of the number combinations or of the spelling of words.

"Arrange the following events in the order in which they occurred" says one author of a history test. Since there is but one order of occurrence and that order is a matter of record about which there can be no dispute, this question is likewise one whose difficulty can easily be determined. Adopting one of the devices of the Army Test whereby a number of ends of sentences are given with instructions to the subject to underscore the right one, another author gives questions of this type: "The Huguenots first settled in *Baltimore, Boston, Charleston, New York, Philadelphia*" Or again, "The Patroons were Dutch *fishermen, fur-traders, land owners, miners, preachers, teachers.*" Again there is only one correct response and its correctness is not questioned by qualified persons. Many other types of questions might be listed—questions the responses to which are either undoubtedly right or undoubtedly wrong. Asking the names of men to be connected with events; requiring the selection, from a given list, of the policies or events favored by a given party; the association of events with dates, of causes with effects, of men with policies,

with principles, or with each other; the classification of men or events in various accepted ways—these are some of the types of exercises now being used. As to the form which answers may take, there is also an increasing richness and variety. One of the earlier devices was to print a question in a box and require the answer to be written in it. Numbers are used to indicate facts, particularly facts regarding the sequence or importance of events. Instead of requiring the pupil to write out the names of men, events, measures, policies, etc., a number of these are printed and check marks are required. The true-false type of test has been introduced.¹ It consists of a series of statements, some of which are true and some of which are false and the pupil is required to indicate the truth or falsity of each statement. The completion type of test has been with us a long time and has proved exceedingly useful.

But in order to be certain that responses to questions are either right or wrong, the questions must, clearly, be questions of fact. One may admit the ease with which such questions may be rated and the cleverness of authors in arranging fact tests and yet remain in doubt as to the value of the results of such tests as measures of ability in history or of the efficiency of the teaching of history. The problem then resolves itself into this: granted that fact questions are easy to devise and easy to rate, what meaning can we give to the results? The critics of measurement have all along been saying that the measurers were formalizing education, that only the least valuable subjects or parts of subjects could be measured, in short, that the only appeal was to the memory or to a few simple skills such as those involved in handwriting.

There are those who seem to think that memory was given us as a human attribute only to be despised. In the vocabulary of the modern pedagogical writer, the word memory is only used in a deprecatory sense. In my own judgment, the case against memory has been vastly overstated. This opinion, however, will not be urged in this connection. But it will be apparent as we proceed that, even when we think we are appealing to a supposedly higher process, we may really be dealing only with a somewhat higher form of memory.

¹ McCall, W. A. "A new kind of school examination," *Journal of Educational Research*, 1:33-46, January, 1920.

Fact questions are those questions whose correct answers are facts. Observe, that the determination of whether a question is a fact question or not depends upon the character of the answer which is accepted as satisfactory. We may ask a child what he *thinks* were the results of the War of 1812; but if we only credit him with a satisfactory answer when he tells the actual results of the war, we are dealing with a fact question, no matter how we delude him or ourselves with the notion that he is fancy-free, because we have asked him what he *thinks*. A fact question appeals to the memory. It may be a date or a name, an incident or an event, a policy or a movement, a cause or an effect, an evaluation or a relationship. It may be bald and meaningless or it may be significant, and even surcharged with emotion.

It appears, therefore, that these fact questions—however rich they may be in meaning—are nevertheless capable of being objectively rated and easily scored.

If, therefore, it can be established that test scores on fact questions are directly related to the thing we call historical ability, we shall be justified in measuring historical ability by means of the ability to answer fact questions. This relationship may be mathematically established if we can secure from the same group of pupils two sets of measures: (1) scores on fact questions and (2) scores in historical ability. I am not sure that I know what historical ability is. Doubtless, psychologically speaking, there is no such thing as historical ability but only many different historical abilities; in other words, the ability is probably plural rather than singular. I presume, however, that we shall come somewhere near the conception of historical ability if we secure measures of ability to think correctly and to judge characters correctly in the field of history. At least in the most elaborate and carefully analyzed history test, namely that of Van Wagenen,² the assumption appears to be made that historical ability is rather fully covered when we have taken into account information, thought, and character judgment.

² Van Wagenen, M. J. *Historical information and judgment in pupils of elementary schools*. New York: Teachers College, Columbia University, 1919.

Van Wagenen, M. J. *American History Scales* (Information Scale, Series A and B; Thought Scale, Series A and B; and Character Judgment Scale, Series A, B, and L). Published by Teachers College, Columbia, University, New York, 1919.

Granted that this is true, two problems challenge our attention: First, the extent to which ability to think and to judge character in the field of history depends upon a knowledge of facts; and second, the extent to which thought and character-judgment questions (using those of Van Wagenen as typical) are really fact questions.

In 1915 I submitted certain fact questions and certain thought questions to some elementary school children in New York. A little more than a year later I gave a different series of fact and thought questions to elementary and high-school pupils at Madison, Wisconsin. I found a correspondence between the ability to answer these fact and thought questions which may be mathematically expressed by a correlation coefficient of about $+0.4$.³ I also found that if one child scores a unit higher than another on the fact questions, he would also, on the average, score 0.89 of a unit higher on the thought questions. From these facts, I derived an equation which would permit the probable score in the thought questions to be forecast from a known score in the fact questions with the chances even that the prediction would not vary more than two out of a possible twenty points from the score which would be obtained by actually administering the thought questions.

The far more extensive testing of Van Wagenen yielded even higher correlations. Between his fact scores and thought scores he found a correlation of $+0.80$ and between his fact scores and character judgment scores the correlation was $+0.78$.

The correlations reported by Van Wagenen are for seventh and eighth grades. I gave his tests to pupils in two high schools, one at Champaign, Illinois, and the other at Danville, Illinois, and obtained correlations between fact and thought scores of $+0.53$ and $+0.60$ respectively. Between fact and character

³ The reader who is unfamiliar with the correlation coefficient may get a rough and ready idea of its meaning by considering that if two series of measures—e.g., scores in fact questions and scores in thought questions—are perfectly and directly related, the correlation coefficient is $+1.00$. Under these circumstances the larger one measure is the larger the other will be. If the measures are not related to each other at all, the correlation coefficient is zero. Coefficients between 0 and $+1.00$ indicate closer direct relationship the more nearly they approach $+1.00$. A coefficient of $+0.4$ indicates a well-marked tendency for the two measures in question to vary together; but it also indicates a rather large number of exceptions to the tendency.

judgment scores the correlations were $+0.45$ and $+0.43$ respectively.

As between fact and thought scores, the relation, therefore, seems to be very close. My high-school results are not as high as Van Wagenen's seventh and eighth grade results because the thought test (a test which Van Wagenen used in grades as low as the fourth) proved to be too easy to distinguish thought ability among the more capable children. But if in grades for which these fact and thought questions are appropriate the correlation is $+0.80$, we are justified in saying that the two tests tell much the same story—in other words, that, after having given the Van Wagenen information test, we shall gain little by giving his thought test.

Nothing, therefore, either in Van Wagenen's testing or mine justifies one in attempting to measure all the desirable outcomes of history teaching. Fact scores seem to be surprisingly sufficient. Moreover, thought and character judgment are hard to rate; and one rater cannot be altogether consistent either with another or with himself at another time. Accordingly, any added contribution which the testing of such abilities may make to our estimate of "general ability" in history will not only be small but will be measured with uncertainty.

To be sure, thought and character judgment do not represent all the desirable outcomes of history teaching. We wish our pupils to grasp the meaning and significance of historical situations, to apprehend changing standards of life, to judge the reliability of historical material or evidence, to distinguish between statements of fact and statements of opinion. This list of aims is taken almost verbatim from Van Wagenen; and doubtless it could be extended. But as to these, may they not all be shown to rest upon a knowledge of facts? How may one "grasp the meaning and significance of historical situations" unless one has at hand the facts about the situations and the additional facts which will give them "meaning and significance"? How may one "apprehend the changing standards of life" unless one knows what the standards were before and after the changes took place? How may one "judge the reliability of historical material or evidence" unless one knows the material and the evidence in question and the facts or criteria on which their

reliability depends? Finally, how may one "distinguish between statements of fact and statements of opinion"—except so far as the form of the statement itself furnishes the distinction—without knowing the facts in the case?

These queries lead to the statement that a knowledge of facts—information, as Van Wagenen calls it—is symptomatic of higher abilities in history. The evidence supporting this view arises not only from the fact, as shown above, that pupils who know facts will also think and judge well but also from two further conditions. First, we may extend the term "facts" to include exact information of any nature in the field of history—information not only as to dates, persons, and places but also as to causation, relationship, political and social movements, economic developments, religious beliefs, wars and rumors of wars, peace and the arts of peace. Second, knowledge of facts conditions higher abilities because many items usually included under these higher abilities are in reality essentially factual. Since Van Wagenen is the first test maker to present a thoroughly worked out scheme of items other than those of information, we may profitably consider some of them. I shall examine the first four questions of his Thought Scale, Series A, together with his acceptable answers to those questions.

The first reads as follows: "Before the steamboats were made people used to travel on the ocean in sail boats. Steamboats were not made until a long, long time after the European people came to make their homes in America. How do you think these early European settlers came to America?" The acceptable answer is "in sail boats" and it is a fact. Therefore the question is a fact question although introduced by the words, "How do you think."

The second thought question reads: "A little before the year 1500 the people of Europe were anxious to find a new way to get to India. Some people thought that India might be reached by sailing westward across the Atlantic Ocean. Columbus was one of these people. It was at this time that Columbus found America. What do you think Columbus was looking for when he found America?" Full credit is allowed for any of the following responses: "A short route to India"; "India"; "Western passage to India"; "Western route to India"; "Northwest passage

to India;" "New way to India;" "A way to India;" "An easy way to India." Here again the answers accepted are facts, and the question is, therefore, a simple fact question. The child is asked what he *thinks* Columbus was looking for, but evidently no thinking is expected.

The third question on the Thought Scale reads: "A hundred years ago it took a letter several days to go from New York to Boston. Today it takes only a few hours. Why do you think it took letters so much longer to go to Boston 100 years ago than it does today?" These answers are allowed full credit: "A hundred years ago there were no railroads;" "no trains or steamships;" "travel by stage coach or horses." Upon what, then, does the ability to answer the question correctly depend? Does it not clearly depend upon a knowledge of facts?

Question four: "The Northmen probably came to America as early as the year 1000, nearly 500 years before Columbus and the Cabots sailed from Europe. There is no record of any one else having come to America before the year 1000. By whom do you think America was first discovered?" Acceptable answers are: "The Northmen," "Vikings," and "Lief Ericcson." If we ask what enables a child to answer "Vikings" or "Lief Ericcson" are we not obliged to say that it is a knowledge of facts? The thought element in all these questions is only in the *form*—a form which employs the word "think."

Not all of the questions on the Van Wagenen Thought Test are so obviously factual, but most of them are; and some of those which are not, call for thinking which has no necessary reference to history.

I shall not examine in detail the questions in character judgment. Each consists first of a narrative and then of a requirement to underline three of ten adjectives which best describe some character or group of characters in the narrative. Reading and language ability are necessary and almost sufficient for success. Facts play no part except as they are given in the narratives; and no other products of instruction in history appear to be required.

Whatever, therefore, may be the desirable outcomes of the teaching of history, it does not appear to be necessary that a valid test should require reactions with reference to all of them.

It may merely require reactions which are symptomatic of them. In other words, we may distinguish between the content and method of teaching on the one hand, and the content and method of testing on the other. Many do not believe this; they regard tests as teaching devices, or at least as encouraging certain types of teaching. This is not always true, and may never be true in a content subject such as history. A test, to be a valid measuring instrument, must provide a reliable quantitative statement which varies in direct proportion to the ability or abilities which we seek to measure. It may be one of the complex of abilities. It may—theoretically at least—be something which is not a historical ability at all just as the expansion of a liquid in a column is not temperature. If it is supposed that by setting up fact tests we encourage formal fact teaching, this supposition may itself be in error. The fact test can be as meaningful as we desire to make it. As I have pointed out, our facts may far transcend those of dates, persons, and places. Again, it will be found that the mere teaching of formal facts will do little to enable a child to pass a well-constructed fact test in history because the right kind of facts are not learned by a dead lift of the memory but by discussion and reflection, by collateral reading and self-expression.

If, after giving a fact test, we should desire to know how capable the pupils probably would be in other ways, we may infer this with substantial accuracy *without giving any other test*. Suppose, for example, that an instructor has given the Van Wagenen information test and wishes to infer the composite ability of the pupils to think and to judge character. The teacher may do this provided the relationship has been established between scores in the information test and the combined scores in the thought and character judgment tests. The correlation coefficient expresses a relationship but not in usable form for this purpose. As a result of the testing carried on in the high schools at Champaign and Danville, we derived an equation by which the required inference may be made. This equation was $y = 72.8 + 0.9x$.⁴ A number of applications of this equation were made and most of them yielded results in close conformity with the actual test

⁴ Such an equation is called a *regression equation*. In this case, x stands for the fact scores and y for the combined scores in thought and character judgment.

results. The first pupil in our alphabetical arrangement had a score in the information test of 33. Substituting this for x in the equation we find that $y=103.5$. This then is the inferred score in thought and character judgment combined. The real score was exactly 100.

The prophesy, however, does not always work as well as this. The second pupil in the alphabetical arrangement scored 36 on the information test. When this value is substituted for x in the equation, it yields a theoretical score of 105.2. The real score was 114. Thus the prophesy is 8.8 units less than the actuality.

It will be interesting to have a general expression for the accuracy that a teacher may expect from the use of this equation. The equation only purports to give the score that on the average would be obtained by giving the thought and character judgment tests a large number of times. In any individual instance the chances are found to be even that the score obtained from the equation will differ from the score obtained by actual testing by not more than eight units. Since there are 144 units in the combined thought and character judgment tests, the chances are one to one (or we may say an even bet) that the score obtained from the equation will be within about 5 percent of the true result.

Thus we have given data and inferences to support the possibility of indirect measurement in history. The doctrine of indirect measurement requires us to find a magnitude which is easy to measure and which also varies closely with the more complex magnitude which we wish to measure but to which it is difficult to apply our instruments. It is maintained that fact questions—when the term fact is given its full and proper meaning—afford an easy means of measurement because of the unequivocal character of the responses and that these responses vary reasonably closely with what we are thus far able to set up as higher abilities in history. We have pointed out that the most notable attempt to measure two of these higher abilities has not succeeded in getting far away from a fact basis. It is also urged that a distinction be made between methods of teaching and methods of test making; and that if the type of testing which we adopt is thoroughly significant in the sense that it varies with the higher abilities, we shall be encouraging the training of these higher abilities when we use it.

THE SELF-SURVEY IN THE ELEMENTARY SCHOOL: AN EXPERIMENT IN COOPERATION

R. H. LANE

Assistant Superintendent of Schools, Los Angeles, California

The following notice appeared in the Los Angeles, California *School Journal* for January 20, 1919:

The Division of Educational Research has in mind for the near future the development of a plan for the Self-Survey of the Elementary School by the Principal and his Associate teachers. This will be accomplished partly by means of standard and other tests, partly by an investigation along uniform lines into several phases of the administration and organization of the school. It is a matter of great gratification to the members of this Division that the courses along this line at the State Normal School, have been so well attended, showing that the time is ripe for an enterprise of this kind.

It must be obvious that such a scheme can be successfully worked out only through the cooperation of those principals who wish to undertake the work. All principals who are interested are asked to communicate at once with the Director, stating that they will be willing to take the time during the spring term to give certain tests and undertake certain studies into the organization and administration of their schools. This will necessitate a few meetings in the near future, but every effort will be made to avoid taking an undue amount of the principals' time. It is suggested that not more than twenty principals form the first group to undertake this work.

In spite of the restriction placed upon the number of members, the class which began work in February was made up of twenty-eight elementary school principals, representing all types of schools from the large well-graded city school in the metropolitan district to the small ungraded rural school. As a check upon the accuracy of the methods employed by members of the class, the director carried on the work outlined for members of the class simultaneously in five other elementary schools, whose principals and teachers contributed largely to the success of this enterprise by their hearty cooperation.

The members of the class were given, in addition to the preliminary instructions for giving the first tests, some general advice as to the spirit in which the enterprise was to be carried out, in something like the following vein:

Before embarking on this enterprise the principal should endeavor to secure to the fullest possible degree the interest and cooperation of his teachers. The tests should not be suddenly sprung upon the teachers without warning, but every effort should be made to prepare carefully and gradually the plans for the campaign. It is inevitable that teachers unfamiliar with educational measurement should look upon a series of tests as an attempt to uncover the weak spots in their teaching and the principal must see to it that his teachers cultivate that fine impersonal attitude which welcomes measurement as a possible solution of certain vexed classroom problems. He should emphasize the fact that one of the greatest contributions of educational psychology is its insistence on the complexity of the learning process, that a child's success or failure in his school work is the product of many factors of which the quality of teaching is only one factor, and that a teacher must not charge herself with the failure of a pupil until she has eliminated all other possible factors from the situation. The reputation of a successful teacher should be such an assured fact both in her own mind and that of her fellow teachers, that the failure of a class in any test or series of tests should not even suggest that the teacher is personally to blame for the result. The finest quality which a teacher can possess is that scientific habit of mind which welcomes occasional class or individual failure as a challenge to her powers of analysis, and interpretation.

Some very definite directions were given as to the preliminary reading each member of the class was to do. Reference was made to the standard books on educational measurement by Thorndike, Monroe, Rugg, and Alexander; to the monographs of the University of Chicago, and to such bulletins as those prepared by the University of Illinois, the city of Boston, the Kansas State Normal School, and Teachers' College, Columbia University.

After the first test had been given by members of the class in their respective schools, the class was called together again for discussion upon the problems which arose in the conduct of the work. One of the chief problems was for the novice to learn the technic of educational statistics. In many cases individual help as well as class help became necessary. Reference was made to the excellent chapters in Rugg's *Statistical Methods Applied to Education*.¹ Finally it became necessary to send out some brief lessons in mimeographed form upon the meaning and method of finding the median, mean, mode, etc.

Seven distinct tests in arithmetic, three tests in reading, the Trabue scale in language, and selections from the Hahn-Lackey scale in geography comprised the material used during the term.

¹ Rugg, H. O. *Statistical methods applied to education*. Boston: Houghton Mifflin Co., 1917.

An effort was made to keep constantly before the class the need for intelligent interpretation of the results in each test. It may be safely laid down as a general rule that any test which does not lead to definite improvement in the progress of pupils is a test which is not worth giving. It is true that having found the "average" or central tendency of a class in some performance the principal is possessed of information which justifies him in saying that a particular class in a particular performance reached, exceeded, or fell below the accepted standard in such case made and provided, but the classroom teacher will not derive much help thereby. In any test four steps should be recognized: (1) The test gives us the facts in the case; (2) It shows which pupils are failing in that phase of instruction covered by the test. Hence if the test is to function: (3) The reason for those individual failures must be found; (4) Remedial measures must be taken to correct those failures.

For a full treatment of this subject of interpretation see Dr. Monroe's *Measuring the Results of Teaching*.² In addition, the results from experiments in the control schools along this line will be found illuminating.

One of the fifth-grade classes tested early in March, attained a median in the Quick Test in Arithmetic³ of 2.5 problems correct in six minutes, an obviously unsatisfactory result. The common prescription of the classroom teacher in such cases is "more drill," but it was pointed out that more drill would be ineffective unless the cause for each individual failure was located. This was done and the following instances of failure noted.

1. In addition several children knew single combinations perfectly, i.e., 6 plus 7, but failed on similar combinations when extended into decades, i.e., 36 plus 7, 47 plus 6, etc.

2. In addition of four columns children had not been taught to place the "carrying figure" above or below successive columns, hence if their attention wavered slightly, the carrying figure was forgotten and the addition had to be gone over from the beginning.

3. In long division many children did not know how to find the first figure of the quotient.

² Monroe, Walter S. *Measuring the results of teaching*. Boston: Houghton Mifflin Co., 1918.

³ *Second year book of the Division of Educational Research, Los Angeles, California, 1919* (Los Angeles City School District No. 21) p. 3.

4. In the following problem a girl worked as follows:

$$\begin{array}{r} 4 \\ 36 \overline{)16236} \\ \underline{144} \end{array}$$

She then asked the observer if one added or subtracted 144 to or from 162!

By means of diagnosis of this kind and very intelligent "follow up" work by the teacher, the class in June attained a median of 5.2 problems correct, an improvement of over 100 percent.

In another control school several children were located who were very poor in oral reading with respect both to speed and accuracy. Careful records were made of these children by means of Gray's Oral Reading Test (Standardized Paragraphs) and individual attention was paid to each child by his teacher with a view to eliminating sources of difficulty. At that time, the writer made the following recommendations to the teachers.

TO THE TEACHER

I present herewith an accurate record of the scores of those children who had unusual difficulty in their oral reading tests. Let us repeat the test in June to note improvement. Failure in oral reading is due to one or all of the following causes: (a) Deficient vocabulary, i.e., lack of a sufficiently large number of words which the pupil is able to recognize at sight; (b) lack of phonic power to analyze unfamiliar words; (c) lack of ability to move the eye along swiftly enough to catch the text in large units or phrases—"eye sweep"; (d) faulty or slovenly enunciation; (e) unfamiliarity in the middle grades with the derivation of words, i.e., word analysis; (f) lack of adequate material for rapid oral reading, i.e., easy supplementary readers.

I therefore recommend exercises as follows:

- (a) Vocabulary drills to enlarge the vocabulary of sight words.
- (b) Phonic drills to give power over unfamiliar words (see Gordon and Beacon Manuals).
- (c) Word analysis drills (see Horace Mann Readers).
- (d) Frequent time exercises in oral reading to cultivate rapidity.
- (e) Enunciation drills.

A repetition of the test in June showed marked improvement.

A word of caution must be given to the principal who gives a test and finds the result below standard. The obvious explanation is that the teacher of the class is a poor teacher, but one soon learns to distrust this easy and obvious explanation. In one of the control schools several primary classes failed in a simple reading test. The personnel of the teaching force was of so high a

character that the factor of poor teaching was quickly eliminated from the situation. After a careful investigation by the principal the following reasons for the failure were found:

1. The pupils in the failing classes had been victims of half-day sessions for several terms.

2. The pupils had had a succession of teachers in the preceding terms, one class alone having had seven different teachers in the ten months of school.

3. There had been several classes in each room at a time, i.e., B2, A2, B3, A3, with the result that in the effort to care for all four classes the amount of time available to the teacher for individual instruction was practically nil.

4. The teachers had not had clearly in mind what constitutes desirable standards of promotion from one class to another, with the result that the classes were poorly graded.

5. Each class had a relatively large number of mentally defective children who were incapable of profiting by classroom instruction in large groups.

At the opening of the fall term 1919, principals were invited to join new Self-Survey Classes along the lines described above. Twenty-two of the original twenty-eight principals registered for membership in an advanced class, while fifty-four other principals registered for elementary classes, necessitating the formation of four sections in the Self-Survey Class, three for beginners and one for advanced students. The work of the beginners for 1919-1920 will be to cover the better known educational tests, while the advanced pupils will study especially group intelligence tests.

Considerable attention has already been given to the Group Point Scale of Intelligence "Schedule D" devised by Dr. and Mrs. Pressey of Indiana University.⁴ This is being given to a large number of A6 classes with some very interesting results. Later on, it is planned to experiment with the Otis Group Tests in eighth-grade classes. The advanced section of the Self-Survey Class will experiment also with Dr. Strayer's Score Card for City School Buildings. Los Angeles has many types of school buildings, and it is believed that an intensive study of the construction of these will result in an awakened interest in better buildings.

⁴ Pressey, S. L. and L. W. "A group point scale of intelligence," *Journal of Applied Psychology*, 2:250-69, September, 1918. The test may be purchased from Indiana University, Bloomington, Indiana.

ANALYSIS OF AND TESTING IN COMMON FRACTIONS

ARTHUR W. KALLOM

Department of Educational Investigation and Measurement, Boston, Massachusetts

In spite of all that has been written in regard to simplicity of work in common fractions, the latest books still continue to give examples involving large denominators and, in the operations of addition and subtraction, continue to use fractions which involve a common denominator so large that the pupils cannot use it to advantage.

In Boston, by the terms of the new course of study, teachers have been asked to restrict their work with fractions to those having a denominator not larger than sixteen and a common denominator not larger than one hundred, and to eliminate all unusual fractions, such as elevenths and thirteenths. This was later emphasized by a test in the four operations prepared by the Department of Educational Investigation and Measurement.

ORGANIZATION OF WORK IN FRACTIONS

Organized work in the study of common fractions was instituted in Boston in December of 1914 by the appointment of a committee of teachers and masters. This committee cooperated with the Department of Educational Investigation and Measurement of the city in the hope of organizing a standardized test in the four operations in fractions. Several months were devoted to the problem, most of the time being consumed in the study of addition and subtraction. In its study the committee distinguished fourteen distinct types of examples in addition and seven in subtraction. The following analysis will show that the classification into the fourteen types in addition is based on two factors, the common denominator and the form of the answer.

ANALYSIS OF TYPE EXAMPLES IN ADDITION OF TWO FRACTIONS

Groups of two or more fractions may be divided into two classes, similar and dissimilar. Similar fractions are defined as fractions having the same denominator. Dissimilar fractions are those which have different denominators. Each one of these

classes may be further subdivided according to whether the result obtained from adding is non-reducible or reducible. If the result is non-reducible, the answer is in an acceptable form after the numerators of the similar fractions have been added. If the result is reducible, it may be changed to a mixed number, to lowest terms, or to both, before the answer can be said to be in an acceptable form. A discussion of the fourteen types is given here because it is believed that the recognition of them will be helpful to teachers in instructing children in the addition of fractions. The reader should bear in mind that these types occur in the addition of two fractions and that the addition of two mixed numbers or a mixed number and a fraction probably creates other types as do examples having three or more addends.

A. Similar Fractions

Four types are all that are possible when fractions are similar.

Type 1.—After adding, the result is non-reducible.

For example:
$$\frac{2}{5} + \frac{1}{5} = \frac{3}{5}$$

Type 2.—After adding, the result is reducible either to an integer or to a mixed number.

For example:
$$\frac{4}{7} + \frac{4}{7} = \frac{8}{7} = 1\frac{1}{7}$$

Type 3.—After adding, the result is reducible to lowest terms.

For example:
$$\frac{5}{9} + \frac{1}{9} = \frac{6}{9} = \frac{2}{3}$$

Type 4.—After adding, the result is reducible both to lowest terms and to a mixed number.

For example:
$$\frac{5}{6} + \frac{5}{6} = \frac{10}{6} = \frac{5}{3} = 1\frac{2}{3}$$

B. Dissimilar Fractions

Under the heading of dissimilar fractions there are three divisions depending on the common denominator.

- a) When the least common denominator is the denominator of one of the fractions.
- b) When the least common denominator is the product of the denominators.

c) When the least common denominator is found by factoring. These will be considered in order.

a) When the least common denominator is the denominator of one of the fractions, there are four types, one non-reducible and three reducible, similar to the preceding four types.

Type 5.—The denominator of one of the fractions is the least common denominator. After adding, the result is non-reducible.

For example:
$$\frac{1}{2} + \frac{3}{8} = \frac{7}{8}$$

Type 6.—The denominator of one of the fractions is the least common denominator. After adding, the result is reducible to a mixed number.

For example:
$$\frac{2}{5} + \frac{7}{10} = \frac{11}{10} = 1\frac{1}{10}$$

Type 7.—The denominator of one of the fractions is the least common denominator. After adding, the result is reducible to lowest terms.

For example:
$$\frac{1}{2} + \frac{3}{10} = \frac{8}{10} = \frac{4}{5}$$

Type 8.—The denominator of one of the fractions is the least common denominator. After adding, the result is reducible to lowest terms and to a mixed number.

For example:
$$\frac{3}{4} + \frac{5}{12} = \frac{14}{12} = \frac{7}{6} = 1\frac{1}{6}$$

b) When the least common denominator is the product of the denominators, there are only two types, one non-reducible and the other reducible to a mixed number. After reducing fractions of these two types to similar fractions and adding the numerators, the result is always a fraction in which the numerator and denominator are prime to each other. This makes it impossible to obtain a result which can be reduced to lowest terms.

Type 9.—The denominators are prime to each other, hence the common denominator must be the product of the denominators. After adding, the result is non-reducible.

For example:
$$\frac{2}{5} + \frac{3}{8} = \frac{13}{40}$$

Type 10.—The denominators are prime to each other, hence the common denominator must be the product of the denominators. After adding, the result is reducible to a mixed number.

For example:
$$\frac{7}{9} + \frac{1}{4} = \frac{37}{36} = 1\frac{1}{36}$$

c) When the least common denominator is found by factoring, there are four types, one non-reducible and three reducible.

Type 11.—The least common denominator is found by factoring. After adding the result is non-reducible.

For example:
$$\frac{1}{6} + \frac{2}{9} = \frac{7}{18}$$

Type 12.—The least common denominator is found by factoring. After adding, the result is reducible to a mixed number.

For example:
$$\frac{1}{4} + \frac{5}{6} = \frac{13}{12} = 1\frac{1}{12}$$

Type 13.—The least common denominator is found by factoring. After adding, the result is reducible to lowest terms.

For example:
$$\frac{1}{6} + \frac{2}{15} = \frac{9}{30} = \frac{3}{10}$$

Type 14.—The least common denominator is found by factoring. After adding, the result is reducible to lowest terms and to a mixed number.

For example:
$$\frac{1}{6} + \frac{9}{10} = \frac{32}{30} = \frac{16}{15} = 1\frac{1}{15}$$

ANALYSIS OF TYPE EXAMPLES IN SUBTRACTION OF FRACTIONS

There are seven distinct types when one fraction is subtracted from another. Owing to lack of space the analysis for subtraction of fractions is not given, but it is on the same basis as has been shown in addition of fractions. The reason for the smaller number of types is because of the impossibility of subtracting one fraction from another and obtaining an integer or mixed number for an answer. Thus there are no subtraction types analogous to Types 2, 4, 6, 8, 10, 12, and 14 in addition. Besides the seven types there are, of course, the subtraction of a mixed number from a whole number or a mixed number from a mixed number.

ANALYSIS OF TYPES IN MULTIPLICATION AND
DIVISION OF FRACTIONS

When the committee approached the study of multiplication and division of fractions, they did not deem it wise to make so extended an analysis. In these two processes the separation into types depended upon the character of the multiplier and multiplicand or the dividend and divisor. Reducible and non-reducible cases were not distinguished. The analysis of multiplication was as follows:

Multiplication of:

- Fraction by a whole number
- Mixed number by a whole number
- Whole number by a fraction
- Whole number by a mixed number
- Fraction by a fraction
- Mixed number by a fraction
- Mixed number by a mixed number

The multiplication of a fraction by a mixed number was not included because there is little or no use for it in adult practice.

In division of fractions the process is either one of two procedures. In the first case, one proceeds to find how many times one number of a certain denomination is contained in a larger number of the same denomination. That is, this process is to determine how many measures of a certain size there are in a measure of a different size. This type of procedure has been called *division by measuring*. In the second case, one proceeds to separate a number into a certain number of parts and to find out the size of each part. This type of division is called *division by parting*. These two cases, measuring and parting, became the basis upon which the analysis in division of fractions was made. In view of these conditions, the following types were selected. Division of:

- Fraction by a whole number (parting)
- Mixed number by a whole number (parting)
- Whole number by a fraction (measuring)
- Whole number by a mixed number (measuring)
- Fraction by a fraction (either)
- Mixed number by a fraction (measuring)

PRELIMINARY WORK

A part of the preliminary work of the teacher committee, prior to any testing, consisted in devising a diagnostic sheet in addition for the use of the teachers. It consisted of 12 examples, one each of Types 3 to 14 inclusive. This sheet of 12 examples was called a form. In order to give the teacher plenty of material to work with, ten different forms were organized. Figure 1 represents Form 1.

FIGURE 1. DIAGNOSTIC SHEET FOR ADDITION OF FRACTIONS
FORM 1

Score.....Time.....Name.....

| | | | |
|--------------------|-------------------|-------------------|-------------------|
| (A) $\frac{3}{16}$ | (B) $\frac{5}{6}$ | (C) $\frac{2}{3}$ | (D) $\frac{3}{4}$ |
| $\frac{7}{16}$ | $\frac{5}{6}$ | $\frac{4}{15}$ | $\frac{7}{8}$ |
| — | — | — | — |
| (E) $\frac{1}{2}$ | (F) $\frac{6}{7}$ | (G) $\frac{2}{9}$ | (H) $\frac{1}{4}$ |
| $\frac{3}{10}$ | $\frac{9}{14}$ | $\frac{7}{10}$ | $\frac{8}{9}$ |
| — | — | — | — |
| (I) $\frac{3}{8}$ | (J) $\frac{7}{9}$ | (K) $\frac{1}{6}$ | (L) $\frac{5}{6}$ |
| $\frac{7}{12}$ | $\frac{8}{15}$ | $\frac{7}{10}$ | $\frac{3}{14}$ |
| — | — | — | — |

The letters at the left of each example indicate the type to which the example belongs. The first example on the sheet lettered "A" represents Type 3, "B" represents Type 4, "C" represents Type 5, etc. On account of their simplicity, Types 1 and 2 were omitted from these forms.

In using the forms the teacher gives each pupil a sheet on which is printed these twelve examples, or twelve others similar to those represented in Figure 1, and requests him to do the

This sheet may be used to preserve the record of each pupil's work. The letters in the first vertical column correspond to the letters on Form 1 (p. 182). After correcting the examples of this form, the teacher or pupil records the results in the column devoted to Form 1. *Those pupils having a perfect score need not waste their time taking any further forms.* Those having wrong work should be given necessary help on the type or types in which they failed: and *after* some drill, they may take a second test represented by the second form. The results are recorded in the column headed "Form 2," and so on. The teacher thus has a continuous record of each pupil. By means of this record a diagnosis of the pupil's difficulties is provided. Effort may be concentrated where it is needed.

Figure 3 shows the record of a girl, M. F., in grade v. A check mark indicates that an example was right, a cross that an example was wrong. The first column shows her result on Form 1. In the succeeding vertical columns is indicated her results on Forms 2, 3, 4, etc. For example, in Form 1, M. F. failed in every type except Type 3 (A). She received individual help and a few days later attempted Form 2, composed of examples similar to those of Form 1. M. F. was then able to do all but Types 8 (F), 9 (G), 13 (K), and 14 (L).

After each form is offered, corrective material is given, laying stress upon the type in which the individual has *failed*. Figure 3 shows that in succeeding tests, although M. F. fails in various types, Type 9 (G) persists as a "trouble maker" until the seventh form has been given. Type 13 (K) is solved correctly on the fourth form, but not again until the tenth form; while Type 14 (L) is solved correctly for the first time at the sixth trial.

Even the ten trials, however, are not enough to fix these two types. Type 12 (J) is "in and out." Of course, only the papers themselves show the trouble in complete detail, but this kind of sheet makes a permanent record which is very valuable to both pupil and teacher. This pupil has made a gain, for she has not only been able to get the examples of each type right at least once, but she has also been able to decrease her time from nine minutes to five, as indicated on the lowest horizontal line.

Thus, the teacher is able by such a selected group of examples, to diagnose the difficulties of her pupils in addition of fractions.

FIGURE 3. RECORD OF M. F. IN ADDITION OF FRACTION

Name M. F.

Grade V

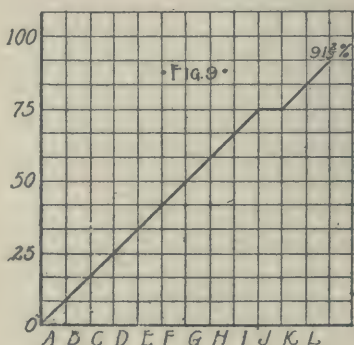
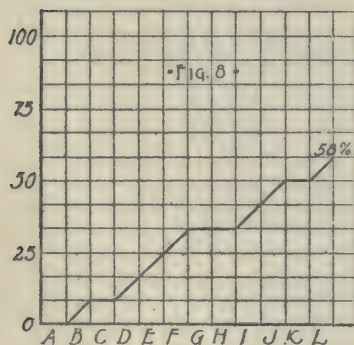
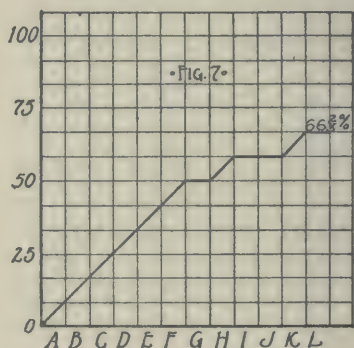
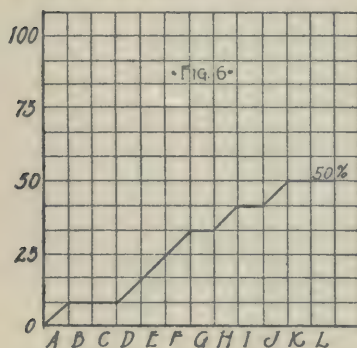
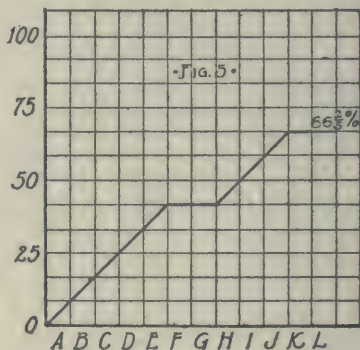
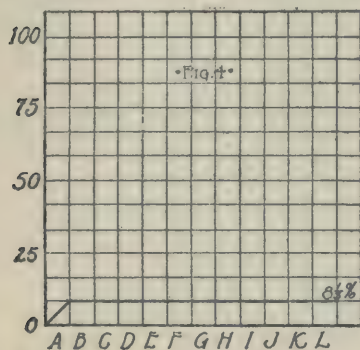
Room 156

| Examples | Form | | | | | | | | | |
|----------|------|----|---|----|----|----|---|---|---|----|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| A | ✓ | ✓ | ✓ | ✓ | ✓ | X | X | X | X | ✓ |
| B | X | ✓ | X | ✓ | ✓ | ✓ | ✓ | ✓ | X | ✓ |
| C | X | ✓ | X | ✓ | ✓ | X | ✓ | ✓ | ✓ | ✓ |
| D | X | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| E | X | ✓ | ✓ | ✓ | ✓ | ✓ | X | X | X | ✓ |
| F | X | X | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| G | X | X | X | X | X | X | ✓ | ✓ | ✓ | ✓ |
| H | X | ✓ | ✓ | ✓ | ✓ | X | X | ✓ | ✓ | ✓ |
| I | X | ✓ | X | X | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| J | X | ✓ | ✓ | X | X | ✓ | X | ✓ | ✓ | X |
| K | X | X | X | ✓ | X | X | X | X | X | ✓ |
| L | X | X | X | X | X | ✓ | X | X | ✓ | ✓ |
| Score | 1 | 8 | 6 | 8 | 8 | 7 | 6 | 8 | 8 | 11 |
| Time | 9 | 10 | 8 | 13 | 13 | 16 | 8 | 8 | 8 | 5 |

Children are growing more and more interested in drawing graphs of their work. This task is simple and comparatively easy when cross-section paper is used. For this work in addition, two lines are drawn at right angles to each other and each twelve spaces long (Figures 4-9). Just below the horizontal line are

placed the twelve letters indicating the types of examples, one in a square. Along the vertical line are shown the percents of the total test, ranging from 0 to 100.

If the example A is right, a short line is drawn from the zero point diagonally upward across the square above A. For each



additional correct example a short line is similarly drawn above the type letter making a continuous curve. In this way, if all examples are right, the curve will be a straight line at an angle of 45 degrees to the horizontal. If an example is wrong a short horizontal line connected with the previous part of the curve is drawn along the base of the appropriate square. Thus the continuous curve will receive no elevation for an incorrect example.

Figures 4 to 9 inclusive show graphically the performance of M. F. on Forms 1, 2, 3, 4, 6, and 10 respectively (see Figure 3). Observe that the end of the curve, when completed, automatically indicates the percent of accuracy for the form to which the curve applies. Thus in the case of M. F., Figure 4 shows that the score on Form 1 was $8\frac{1}{3}$ percent. According to Figure 5, her score on Form 2 was $66\frac{2}{3}$ percent. Figure 9 shows a percent of $91\frac{2}{3}$ for Form 10.

The shape of the curve also indicates, in some measure, the way the pupil learns. The interest of the pupil is held in attempting to make a straight line. The plan, therefore, becomes a real means of motivating the work.

These two methods of keeping a record enable the teacher to determine at once where the pupil is having difficulty, and it must be remembered that a pupil's ability to add fractions is no greater than his ability in each one of these fourteen types. Such records indicate that a pupil may be able to do the simpler types without being able to do the more complicated ones—e.g., Types 13 and 14. In other words, *the general knowledge that a pupil can or cannot add fractions is not sufficient*. On that basis no teacher can accurately diagnose difficulties or intelligently apply remedies. It is only when the teacher takes the various possibilities—possibilities such as this analysis of operations in adding fractions supplies—and tests the children with all of these that she is able to determine whether each pupil will probably be able to add *any* example in fractions with which he may be confronted.

Monroe directs attention to this same fact when he calls "the situation which the teacher faces a complex one." He adds that: "Before she can intelligently direct her efforts as an instructor she must diagnose the situation."¹ This may be done only by taking all the possibilities into consideration.

¹ Monroe, Walter S. and De Voss, J. C. and Kelly, F. J. *Educational tests and measurements*. Boston: Houghton Mifflin Co., 1917, p. 49.

RESULTS OF PRELIMINARY WORK

Forms similar to those in Figure 1 (page 182) were used in the fifth grades of five districts in Boston in the spring of 1915. The results of this testing, as sent to the office of the department, showed that in general the accuracy increased and the time necessary to do the examples shortened as the children worked each successive form. In addition to this, the data gathered from this material showed:

- a) That testing of this kind would be likely to produce results which would show the ability of children to add fractions.
- b) That children might be able to do non-reducible types but would have trouble in doing types where reduction was necessary.
- c) That the time required to do the twelve examples on each form varied greatly. The shortest time recorded by an individual was two minutes and the longest forty-one minutes.

Similar forms are now being printed for use in subtraction, multiplication, and division of fractions.

SUGGESTIONS FOR DIAGNOSIS

It has been pointed out in the foregoing that satisfactory diagnosis can be carried on only when all types are represented. The knowledge that a pupil can multiply $456 \frac{1}{3}$ by 2 is not sufficient, because the example is so constructed that it does not bring in all the elements of a mixed number multiplied by an integer. A pupil may do this type without being able to do a similar example when the product of the fraction by the integer is equal to or more than a whole number.

What is meant, then, by diagnosis? By diagnosis is meant noting certain existing symptoms and finding out from them what is the trouble. This is usually done by correcting the pupil's examples. Each pupil may have a different trouble and, therefore, require a different treatment. Thus the importance of the individual is brought to the front. There may be two reasons for looking over a set of papers: (1) simply to mark examples right or wrong; (2) to gain some knowledge which will tend to correct pupil's work. Diagnosis emphasizes the latter reason.

There are four principles which underlie any diagnosis of the results of a test aside from the necessity of having all types of examples represented.

a) Trouble with individual pupils cannot be discovered by studying and emphasizing the work of the children who have examples right. The principle that success comes through failure must be practiced here, and the teacher will find her richest material for study in the poorest papers.

b) After selecting the papers with low scores it is necessary to find out from each paper, if possible, how the pupil did his work. The fact that a pupil usually does each type of example in the same way makes this possible in a large number of cases. In many cases, however, it is not possible to determine how the work is done because of the queer methods used by different individuals. The following are a few of the methods used by pupils in a test given in December, 1917:

$$(1) \begin{array}{r} 275 \\ \underline{8\frac{3}{4}} \\ 23 = 4 \times 5 + 3 \\ 2200 = 8 \times 275 \\ \hline 22023 \end{array}$$

$$(2) \begin{array}{l} 4\frac{7}{8} \times \frac{1}{8} = 7\frac{1}{2} \\ 4\frac{7}{8} \\ \underline{\frac{1}{8}} \\ 7 = \frac{7}{8} \times \frac{1}{8} = \frac{7}{8} \times \frac{8}{8} = \frac{7}{1} = 7 \\ \underline{\frac{1}{2}} = \frac{1}{8} \times 4 \\ 7\frac{1}{2} \end{array}$$

$$(3) \begin{array}{l} 4\frac{7}{8} \times \frac{1}{8} = \\ \frac{39}{8} \times \frac{1}{8} = 313 \quad (8 \times 39 + 1) \end{array}$$

$$(4) \begin{array}{l} 3\frac{3}{4} \div \frac{1}{5} = 13. \\ 3\frac{3}{4} = 375 \quad (3.75) \\ \frac{1}{5} = 20 \quad (.20) \end{array}$$

$$20 \overline{) 375} \begin{array}{l} 13 \\ 20 \end{array}$$

$$(5) \begin{array}{l} \frac{1}{8} \times 6 = 64 \\ \frac{1}{8} \times 6 = \end{array}$$

$$\begin{array}{l} \frac{1}{8} \times \frac{3}{48} = \frac{3}{192} = 64 \\ 4 \quad (6 \times 8) \end{array} \quad \begin{array}{l} 64 \\ 3 \overline{) 192} \end{array}$$

All the methods illustrated above, with many others, show the almost unbelievable ways used by children even though they have been taught proper methods. If an examination of the paper fails to show how the work was done, then a session with

the pupil is, of course, necessary. What is essential to discover is *whether the pupil has anything to build upon*. If he has little except a collection of bad habits, then there is a long, weary struggle to overcome these habits and establish habits of right thinking and right action before results can be obtained. For in the last analysis accuracy in all mechanical work depends upon the formation of right habits of work.

c) The nature of the corrective material to be given is determined by the results of the study of the papers made under (a) and (b). To try to determine this before the study means wasted time for both teacher and pupil.

d) The most important step of all is the keeping of a record of each pupil's performance. In no other way can one tell whether the pupil is advancing or not.

PLAN OF RECORDING RESULTS

As a result of this study of common fractions, how may a teacher effectively check the work in her grade? How may she determine, and keep in some permanent form, a record of the pupil's ability in the four operations in common fractions? The accompanying form of record (Figure 10) has been used and has proved effective. This sheet may be duplicated so that each pupil may have a copy. It was planned to be used as follows:

The teacher may give examples in addition similar to the types illustrated on page 178 or she may give examples of the various types into which the other three operations have been analyzed.

If the pupils get the answers right, the appropriate types may be checked in the column marked "R," if wrong they may be checked in the column marked "W." The checks in the "wrong" column should be changed as fast as the pupil has mastered the types, or a separate record sheet may be used for each trial. After giving a series of lessons covering the various types and recording results, the teacher has a record of the ability of each individual in the room showing his strength and weakness. If a pupil fails in a verbal problem, it can be determined immediately by reference to this record whether or not the difficulty is in the mechanics of the problem. The mechanics may be wrong, and still the pupil may have shown in previous work that he knew how to do

FIGURE 10. RECORD IN COMMON FRACTIONS OF

Name..... School.....
 Age..... Room.....
 Grade.....

| | R. | W. | | R. | W. |
|---------------------------|----|----|----------------------------|----|----|
| Addition | | | Multiplication | | |
| Type 1..... | | | Fraction by an integer... | | |
| 2..... | | | Integer by a fraction.... | | |
| 3..... | | | Integer by a mixed number | | |
| 4..... | | | Mixed number by an in- | | |
| 5..... | | | teger..... | | |
| 6..... | | | Fraction by a mixed num- | | |
| 7..... | | | ber..... | | |
| 8..... | | | Mixed number by a frac- | | |
| 9..... | | | tion..... | | |
| 10..... | | | Fraction by a fraction.... | | |
| 11..... | | | Mixed number by a mixed | | |
| 12..... | | | number..... | | |
| 13..... | | | | | |
| 14..... | | | | | |
| Mixed numbers..... | | | | | |
| Subtraction | | | Division | | |
| (Without "Borrowing") | | | | | |
| Type 1..... | | | Fraction by an integer... | | |
| 2..... | | | Integer by a fraction.... | | |
| 3..... | | | Integer by a mixed number | | |
| 4..... | | | Mixed number by an in- | | |
| 5..... | | | teger..... | | |
| 6..... | | | Mixed number by a frac- | | |
| 7..... | | | tion..... | | |
| Fraction from mixed num- | | | Fraction by a fraction.... | | |
| ber..... | | | | | |
| Mixed number from mixed | | | | | |
| number (with "bor- | | | | | |
| rowing")..... | | | | | |
| Fraction from integer.... | | | | | |
| Fraction from mixed num- | | | | | |
| ber..... | | | | | |
| Mixed number from integer | | | | | |
| Mixed number from mixed | | | | | |
| number..... | | | | | |

the type of mechanical work involved. If so, the teacher's work should be quite different from what it would be if the pupil did not know the mechanics involved. In any case, the important point is that the teacher can easily tell from her record in what way and to what extent she should help the pupil.

CONCLUSIONS

I. Modern teaching requires an analysis of conditions such that the teacher may know *all* the possibilities in the subject under discussion.

II. Knowledge of the subject demands knowledge of all parts of the subject as determined by the analysis which has been made.

III. Diagnosis of results depends upon four principles:

1. Study the papers of the poorest pupils.
2. Find out how each pupil did his work.
3. Determine from this the kind of corrective material to be given.
4. Keep a record of each pupil's performance.

THE ENCOURAGEMENT OF INDIVIDUAL INSTRUCTION BY MEANS OF STANDARDIZED TESTS

MATTHEW H. WILLING
University School, Cleveland, Ohio

To encourage an intelligent individualization in school instruction is, to the writer's mind, the most pressing duty of modern educational supervision. It is one aspect, and that a decidedly technical one, of the whole movement toward adapting the school to the individual. The general problem has been attacked administratively in reforms relating to promotion, grading, the size of classes, and the nature and extent of the curriculum. But in the teaching process there have been few evidences of an appreciation of the problem. Regardless of improved grading, reduced classes, greater flexibility of promotion, and the like, the teaching technic still remains largely managerial rather than tutorial. The existence of individual differences is readily enough admitted, but they are rarely defined or their extent calculated. Their implications for instruction are pretty widely ignored. Alert supervision will tend to change this situation; will challenge the sanctity of the eternal test recitation; will decry uniform assignments and uniform drill; will insist on analytical marking; will stimulate teachers and pupils alike to determine individual needs and apply individual remedies.

As an instrument in bringing about the desired change in instructional emphasis, the standardized test is helpful. In three ways, at least, it meets initial problems. (1) It furnishes concrete, intelligible evidence of the presence and extent of individual differences within school groups. (2) It defines, in many instances, the exact nature of the differences or makes such definition possible. (3) It permits the establishment of definite standards for individual accomplishment in many fundamental phases of school learning.

Little need be said of the service performed by standardized tests in revealing the variation in ability within school groups. Throughout the now abundant literature on the measurement of achievement in school subjects a foremost place has been given to

the phenomena of individual differences and their implications. It has been stated repeatedly that an average is a feeble guide for instruction when compared with a distribution table, and that for the teacher the most important data from these tests are not averages, or medians, or other records of central tendencies, but rather the individual deviations from these measures of type. Statistical, verbal, and graphical representation of such deviations are the devices to use in making teachers conscious of the most significant facts about their pupils. The great advantages which such representations have when derived from standardized tests are their precision and the fact that they have a common meaning for all who understand them.

Almost from the outset, however, any supervisor who uses the standardized test to induce greater individualization of instruction will find it necessary to supplement the mere statement of individual differences with precise information as to their nature. Except in the most simple kinds of school work a mere score is not particularly illuminating, even though the unit of measure is defined. A standardized test is helpful to the teacher in just about the degree that it is itself diagnostic or makes diagnosis possible. Furthermore, it is helpful in just about the degree that the teacher exploits the diagnostic possibilities herself.

The way to do this, however, must be shown. As an illustration of the treatment that may be given profitably to the papers of a standardized test, devised to permit diagnosis, the following abridged supervisor's report on the accomplishment of a seventh grade is offered. The material is Monroe's Reasoning Test in Arithmetic Test II, Form 1, given January 28, 1919.¹

The median scores made in this test by the two seventh grade classes are as follows:

| Class | Rate Score | Principle Score | Accuracy Score |
|-----------------------------------|------------|-----------------|----------------|
| Mr. E's Class.... | 13.7 | 23.6 | 15.1 |
| Mr. W's Class.... | 11.1 | 19.8 | 15.5 |
| Monroe's Tentative Standards..... | | 21 | 14 |

¹ For detailed description of this test see Walter S. Monroe's *Measuring the Results of Teaching*. A copy of the test can be obtained from the Bureau of Educational Research, University of Illinois, Urbana, Illinois.

One class appears to be about as capable as the other in this test, while neither is at present much above the tentative standard. There is a slight indication that Mr. E's pupils are less accurate than Mr. W's. This apparent difference, however, may be wholly due to the time of day when the test was given to the respective classes. Mr. W's class was tested in the morning at the second period, while Mr. E's was not tested till afternoon at the eighth period. Pupils of this age may be less painstaking and accurate at the close of a long school day than at the beginning.

The sixth grade was given the same test a few days earlier and scored as follows: rate 7.7, principle 15.5, accuracy 10. Pupils in the seventh grade making lower scores than these, clearly need to be made subjects of special attention at once. A list of such is given herewith:² . . .

This test calls for ability in the fundamental operations with integers, common fractions, and decimals. The actual mechanical processes involved are not such as to tax any seventh grade child provided with average preparation. The types of reasoning called for are also completely within the experience of boys and girls of this grade, and no problem demands more than two steps. In the case of only two problems have our pupils any excuse for failure on the plea of lack of training; namely, the eleventh and the thirteenth. The eleventh calls for an understanding of the simplest sort of commercial accounting, and the thirteenth requires a knowledge of the number of pounds in a ton. So little attention has been paid as yet to commercial forms, and so long a time has elapsed since the intensive study of denominate numbers, that many failures in the problems named were to be expected.

The following list of reasons for failing in principle or in accuracy of calculation covers the present situation pretty completely, as revealed by the papers themselves:

ERRORS

- A. In Principle:
 - 1. Faulty reasoning
 - 2. Careless reading
 - 3. Apparent ignorance of terms and concepts used
 - 4. Incorrect statement of division ($\frac{3}{8} \div 6$ instead of $6 \div \frac{3}{8}$), though correct operation
 - 5. Additional operations not called for
- B. In Accuracy:
 - 6. Errors in fundamental operations with integers
 - 7. Errors in manipulating common fractions
 - 8. Errors in placing decimal point
 - 9. Error in copying
 - 10. Skipped problem; probably unable to solve.

² In the original report, names of pupils were given at this point.

Table I contains a brief description of each problem, the number of pupils attempting, the number scoring in principle, the number scoring in accuracy, and a record of the more common errors with the number of pupils (in parentheses) making each error.³

TABLE I. PROBLEM RECORD

| PROBLEM | ATTEMPTS | NUMBER OF PUPILS HAVING | | TYPES OF ERRORS |
|---|----------|-------------------------|----------------|------------------|
| | | Principle Correct | Answer Correct | |
| 1. One step. ($a+b=?$) Addition of fractions | 41 | 39 | 37 | 1(2), 6(2) |
| 2. One step. ($a-b=?$) Subtraction of fractions | 41 | 38 | 27 | 7(7), 9(2), 1(2) |
| 5. Two steps. ($a \div b \times c=?$) Div. and mult. of dec. | 41 | 37 | 35 | 1(4), 6(2) |
| 11. Two steps. ($a+b+ \dots -h=?$) Addition and sub. of dec. | 31 | 7 | 3 | 3(20), 10(10) |

In the original report there followed here a lengthy discussion of each of the ten types of error recognized, with a statement of the number of pupils making each error, together with names in aggravated cases. Then a final table was offered giving the individual records of the pupils on each problem, on types of error, and in the various test scores. This table was presented in the accompanying form (Table II).

³ Only four of the fifteen problems are here shown. Forty-one pupils took the test.

TABLE II. INDIVIDUAL RECORD

| Pupil | | Problems | | | | | | | | | | | | 7-14 omitted | 15 | | Tot Att. | Tot Pr. | Tot Ac. |
|-------|------------|----------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--------------|----|----|-------------|------------|------------|
| | | 1 | | 2 | | 3 | | 4 | | 5 | | 6 | | | | | | | |
| | | Pr. | Ac. | Pr. | Ac. | Pr. | Ac. | Pr. | Ac. | Pr. | Ac. | Pr. | Ac. | | | | | | |
| A | Error..... | 1 | | 1 | | | | 7 | | 1 | | 1 | | 1 | | | | | |
| | Credit.... | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 15 | 3 | 1 | |
| B | Error..... | | | 1 | | | | 1 | | 1 | | | 6 | | | | | | |
| | Credit.... | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 15 | 11 | 9 | |

The reader will observe that the above table is broken between the sixth and fifteenth problems. The Total Attempts, Total Credits in Principle and Total Credits in Accuracy are not therefore the totals of the figures actually shown. The table reads: Pupil A made errors of Type 1 in principle (see code above) on problems 1, 2, 5, 6, and 15. He obtained credit in principle on problems 3 and 4 with one additional credit on a problem not shown, thus making the 3 credits indicated under "Tot Pr." He also obtained a credit in accuracy on problem 3, and on no other, thus making the total credit of "1" indicated in the last column. It is to be noted that if there is an error in principle on any problem no credit for accuracy is possible.

A supervisor speedily discovers that the giving of tests, the demonstration of pupil variation, and the explicit diagnosis of individual weaknesses will not of themselves guarantee any marked change in teaching method. Group methods constitute too much of an obsession to be easily thrown off even in part. It must be made insistently clear that the results desired from teaching are individual results and not class results.

The standardized test performs a distinct service in pointing out where to look for specific troubles when minimum requirements are furnished in terms of the units of the test. For

example, standards in handwriting may be stated in some such form as the one shown in Table III.

TABLE III. STANDARDS IN HANDWRITING

| Grade | Ayres' Quality | Letters per Min. | Percent of Class |
|-------|----------------|------------------|------------------|
| III | 40 | 45 | 80 |
| IV | 40 | 50 | 80 |
| V | 50 | 55 | 80 |
| VI | 50 | 60 | 80 |
| VII | 60 | 65 | 80 |
| VIII | 60 | 70 | 80 |

A similar statement of standards is possible today in practically all elementary-school fundamentals and in certain high-school subjects as well. Increased individualization is sure to follow such prescriptions, for in no other way can the required results be achieved with reasonable economy.

The extensive use of the standardized test for the purposes outlined above, if for no other, is strongly recommended to supervisors. It will do a great deal at once to center attention on the pupil instead of on the class. It will not solve the ultimate problem of teaching technic, but it will produce the necessary antecedent conditions of enlightenment and stimulation. Therefore, when this report was submitted, the teachers were asked to make it the basis not only of a more analytical procedure in terms of the subject, but, especially, of a more individualized procedure in terms of the pupil. Attention to the specific troubles of each of these pupils as revealed in the Individual Record (Table II) was held to be far more profitable than marshaling them as a group through daily recitation maneuvers.

CORRELATION OF PHYSICAL HEALTH AND MENTAL EFFICIENCY

R. L. SANDWICK

Principal, Deerfield-Shields Township High School, Highland Park, Illinois

The following study, made in one school, of the relation between mental power as determined by an intelligence test and comparative freedom from physical defects, is submitted to the JOURNAL OF EDUCATIONAL RESEARCH in the hope that similar studies may be made in other schools. It is important to know whether the remarkable parallelism between mental inefficiency and physical defects here shown is merely accidental, or indicative of a general condition.

The Group Intelligence Test by Rugg and Freeman¹ was given to 423 of the students enrolled in the Deerfield-Shields High School on September 18, 1919. The statistical work in connection with the test was handled by Mr. Harry G. Pertz, who presented, among other matters, a chart showing the scores of the forty ablest students as determined by the test, and of the forty students making the lowest scores. He also arranged the whole 423 students in order of scoring ability from the highest with a score of 81.5 to the lowest with a score of 8.

Mr. Pertz had made a similar gradation of these students a year ago after all of them had taken the Kansas Silent Reading Test. In general the two tests show parallel ability. Most students who ranked high in the Kansas Silent Reading Test also ranked high in the Group Intelligence Test. There were a few conspicuous examples of the failure of this general rule. One student who ranked seventy-ninth in the test last year ranked second this year. It was known that this particular boy had had some rather serious physical defects remedied by surgical operation—defects which had interfered with his health the previous year. He had spent the summer on a farm in the open air, and had gained about thirty pounds in weight in a year's time. Being impressed by

¹ The Freeman-Rugg Group Intelligence Tests were designed to measure the general intelligence of junior and senior high-school pupils. Address, University of Chicago Bookstore, Chicago, Illinois. Price 5 cents a sample set; 3 cents each in quantities

TABLE I. STUDENTS HAVING FORTY HIGHEST SCORES

| Student No. | Intelligence Score | Physical Defects |
|-------------|--------------------|--|
| 1 | 81.5 | No defects |
| 2 | 78.0 | No defects |
| 3 | 75.0 | No defects |
| 4 | 75.0 | Teeth |
| 5 | 73.5 | No defects |
| 6 | 72.5 | No defects |
| 7 | 71.5 | No defects |
| 8 | 71.0 | No defects |
| 9 | 71.0 | No defects |
| 10 | 70.0 | No defects |
| 11 | 70.0 | No defects |
| 12 | 70.0 | Teeth, vision (not bad) |
| 13 | 70.0 | Skin disease (slight) |
| 14 | 69.0 | Teeth |
| 15 | 67.0 | No defects |
| 16 | 66.5 | No defects |
| 17 | 66.5 | No defects |
| 18 | 66.0 | Vision (slight), teeth (not bad), glands (temporary from cold) |
| 19 | 66.0 | Christian Scientist (excused from physical examination) |
| 20 | 65.5 | Pulmonary disease, nasal breathing |
| 21 | 65.5 | Slight acne, hearing (from cold) |
| 22 | 64.5 | Teeth (small filling dropped out) |
| 23 | 64.0 | No defects |
| 24 | 64.0 | No defects |
| 25 | 64.0 | No defects |
| 26 | 64.0 | No defects |
| 27 | 64.0 | Cardiac (result of influenza) |
| 28 | 63.5 | Vision (wears glasses) |
| 29 | 63.0 | No defects |
| 30 | 63.0 | Tonsils (temporarily enlarged) |
| 31 | 62.5 | Vision (eyes slightly crossed) |
| 32 | 62.5 | Cardiac |
| 33 | 62.5 | Tonsils |
| 34 | 62.0 | Teeth (not serious) |
| 35 | 62.0 | Tonsils (from cold), vision (slight astigmatism), glands |
| 36 | 62.0 | No defects |
| 37 | 61.0 | Christian Scientist (excused from physical examination) |
| 38 | 61.0 | Tonsils (slightly enlarged), teeth |
| 39 | 60.5 | No defects |
| 40 | 60.0 | No defects |

TABLE II. STUDENTS HAVING FORTY LOWEST SCORES

| Student No. | Intelligence Score | Physical Defects |
|-------------|--------------------|--|
| 1 | 29.5 | Skin, cardiac, pulmonary |
| 2 | 29.5 | Nervous disease, cardiac, pulmonary, tonsils, adenoids |
| 3 | 29.0 | Enlarged glands, goitre, cardiac disease, skin, vision, teeth, tonsils |
| 4 | 29.0 | Enlarged glands, cardiac, pulmonary, orthopaedic |
| 5 | 29.0 | Enlarged glands, goitre, pulmonary, other defects of eye, tonsils, teeth |
| 6 | 29.0 | Enlarged glands, vision, nasal breathing, tonsils |
| 7 | 29.0 | Cardiac, pulmonary, orthopaedic, vision, tonsils |
| 8 | 28.5 | Glands, vision, nasal breathing, tonsils |
| 9 | 28.5 | Pulmonary, vision, teeth, tonsils |
| 10 | 27.5 | Enlarged glands, cardiac, pulmonary, vision |
| 11 | 27.5 | Enlarged glands, cardiac, pulmonary, vision, teeth, tonsils |
| 12 | 27.0 | Enlarged glands, goitre, tonsils |
| 13 | 27.0 | Tonsils, vision |
| 14 | 27.0 | Pulmonary, teeth |
| 15 | 27.0 | Tonsils, teeth |
| 16 | 27.0 | Teeth, tonsils, adenoids |
| 17 | 26.5 | Teeth, tonsils, adenoids |
| 18 | 25.0 | Glands, skin disease |
| 19 | 25.0 | Glands, goitre, cardiac |
| 20 | 24.5 | Enlarged glands |
| 21 | 24.5 | Tonsils, vision, cardiac |
| 22 | 24.5 | Vision, goitre |
| 23 | 21.0 | Cardiac, teeth, vision, tonsils |
| 24 | 21.0 | Vision, goitre |
| 25 | 20.5 | Cardiac, skin, rachitic type, teeth, tonsils |
| 26 | 20.5 | Pulmonary, tonsils |
| 27 | 20.0 | Cardiac, teeth, tonsils |
| 28 | 20.0 | Teeth, adenoids |
| 29 | 19.0 | Vision |
| 30 | 19.0 | Enlarged glands |
| 31 | 18.0 | Teeth, nasal breathing |
| 32 | 17.0 | Enlarged glands, goitre, cardiac, teeth, tonsils |
| 33 | 15.5 | (Not examined) |
| 34 | 15.0 | Nasal breathing, teeth, glands, tonsils |
| 35 | 14.0 | Teeth, goitre, glands |
| 36 | 13.0 | Cardiac, tonsils |
| 37 | 13.0 | Tonsils, glands |
| 38 | 12.0 | Disease of eye, teeth, tonsils, discharging ear |
| 39 | 10.0 | Cardiac, pulmonary, vision, teeth, hearing |
| 40 | 8.0 | Tonsils |

TABLE III. SUMMARY

| | In Group with Highest Scores | In Group with Lowest Scores |
|--|---------------------------------|--------------------------------|
| Total number of defects..... | 27 | 125 |
| Average number of defects per student... | 0.71 | 3.41 |
| Percent having no defects..... | 52.5 | 0 |

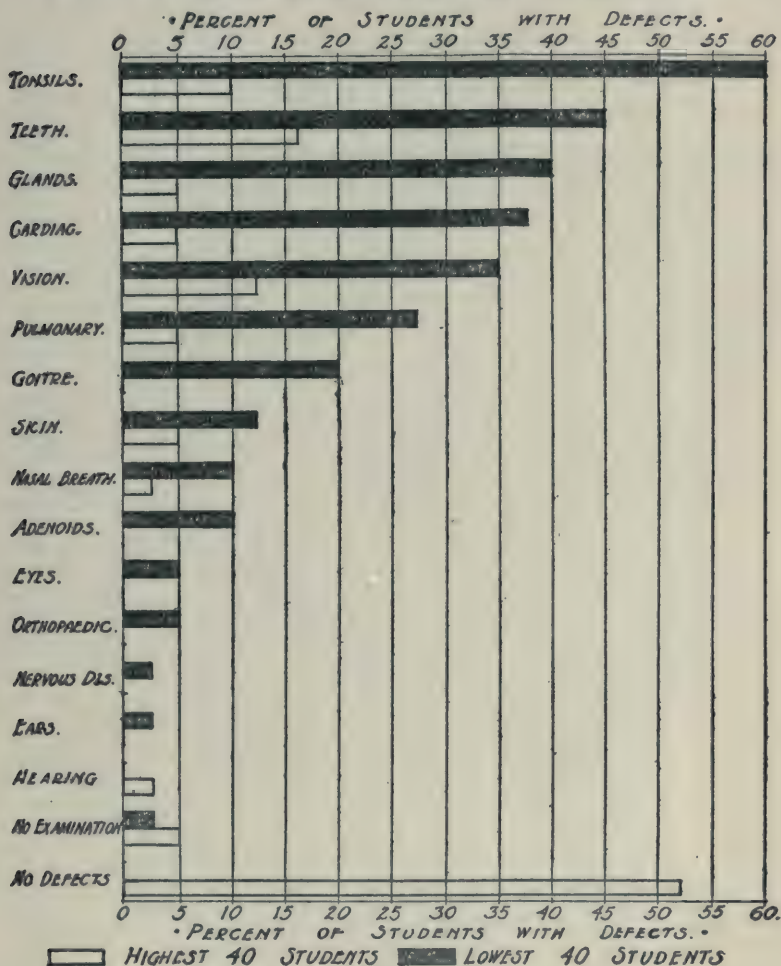
this boy's showing in the test, and thinking that physical condition might bear some relation to defects, we put the list of forty students with best scores and the list of forty students with poorest scores into the hands of our school nurse with instructions to give us the health card of every student. The records are given above.

It was noticeable that the defects among the ablest group were not only fewer in number but also less serious in character. Some of them, indeed, were only temporary. Most of the defects among both groups are entirely remediable.

In Figure 1 details are shown as to the existence of each defect in the groups having highest and lowest scores. Of the fifteen defects only one occurred oftener among the students having the highest scores. This was due to a single case of defective hearing which was found in a girl belonging to the high group. Of course, this case has no significance whatever beside such facts as these: that 60 percent of the low group had defective tonsils compared with 10 percent for the high group; that 45 percent of the low group had defective teeth compared with 17 percent for the high group; etc.

Although the number of cases is small, the investigation adds something to the accumulating evidence that the child of good intellectual ability is also of good physical ability. In other words, it supports the view that the gifted child is not likely to be the physical and nervous wreck that some have supposed. It also contributes to the growing conviction that good abilities may be and often are masked through the presence of physical defects. And it points to the unmistakable duty of the school to provide for every child early and frequent medical inspection with a systematic follow-up in applying remedies.

FIGURE 1. PROPORTION OF DEFECTS AMONG THE FORTY PUPILS HAVING THE HIGHEST AND LOWEST INTELLIGENCE SCORES



A STATE-WIDE SCHOOL-BUILDING SURVEY

GEORGE DRAYTON STRAYER

Teachers' College, Columbia University

State control of school buildings has been exercised in the United States in some degree by having the plans for new buildings passed upon by the state architect or by some special officer in the state department of education. Laws have also been enacted which require inspection and approval by the state department of education when repairs exceeding a minimum amount are undertaken. This control has helped over a period of years to develop a somewhat better type of building as the population has increased and as new buildings have been needed. There has been little or no control, however, over buildings which were already in use and which were not reconstructed so as to come under the provisions of the law.

Throughout the United States there are today thousands of school buildings which are unfit for use, and in many cases dangerous in the extreme because of the unhygienic conditions or fire risks involved. There is need throughout the country for state-wide school-building surveys which will reveal the conditions and which will arouse the public to the need of well-considered school-building programs.

The state of Delaware is the first to undertake this particular type of survey. Through the generosity of Mr. Pierre S. du Pont, the state has been provided with a fund for the construction of its school buildings. In every case the local community is to provide funds equal to or more than those provided by the du Pont fund, except in the case of schools for colored children, where the whole cost will be borne by those administering Mr. du Pont's gift. In order to know definitely the condition of school buildings throughout the state, Mr. du Pont employed the writer and his colleagues to make a school-building survey of all of the school-houses in the state outside of the city of Wilmington.¹ Every

¹ Strayer, George D., Engelhardt, N. L., and Hart, F. W. *General report on school buildings and grounds of Delaware*. (Bulletin of the Service Citizens of Delaware, v. 1, no. 3). Wilmington, Delaware: Service Citizens of Delaware, 1919.

one of these school buildings was visited by one or more of the survey commission and rated with respect to site, building, service systems, classrooms, and special rooms. Each of these main divisions of the score is made up of certain subdivisions appearing upon a score card which permits of a maximum of one thousand points for a perfect building.²

Included under *site* are location, drainage, size and form, and flagpole. Under *building* are included placement, gross structure, and internal structure. Under *service systems* are included heating and ventilation, fire protection, cleaning system, artificial lighting, schedule and emergency equipment, water supply system, and toilet systems. Under *classrooms* are included arrangement, construction and finish, illumination, cloakrooms and wardrobes, and equipment. Under *special rooms* are included rooms for general use, such as playroom, community room, library, and lunch room; officials' consultation room, and other special service rooms (e.g., those for industrial arts and household arts), fuel room, storage room, and the like. Each of these subdivisions is again analyzed so as to indicate the detail which one needs in scoring a school building. Accompanying the score card is a set of standards which has been worked out with the cooperation of a large number of persons interested in educational administration, in school architecture, and in schoolhouse construction.

The astonishing thing revealed by the survey in Delaware is the almost complete lack of suitable buildings and equipment throughout the state. This is particularly true with respect to the village and rural schools—i.e., the schools having one to four teachers. If one accepts four hundred points out of a thousand as the minimum score for a building that is to be permitted to continue in use, there are, of the school buildings of more than four teachers, just seven buildings in the state, outside of Wilmington, which should be left standing. The following table gives the scores for these buildings.

An analysis of the partial scores allowed establishes the fact that these low scores do not depend upon a failure to provide any

² Strayer and Engelhardt Score Card for City School Buildings and Score Card for One to Four-Teacher School Buildings. Published by the Bureau of Publications, Teachers' College, Columbia University, New York City.

one element that goes to make up a satisfactory building, but, rather, that the buildings are unfit for use, regardless of the particular standard applied.

TABLE I. SCORES FOR CITY SCHOOL BUILDINGS OF MORE THAN FOUR TEACHERS

(Arranged in Order of Rank)

| Scores of Buildings | Location |
|---------------------|--|
| 850-899 | Alex I. du Pont |
| 800-849 | Caesar Rodney |
| 750-799 | (None) |
| 700-749 | " |
| 650-699 | " |
| 600-649 | Harrington |
| 550-599 | Bridgeville |
| 500-549 | Greenwood |
| 450-499 | Frederica, Mt. Pleasant |
| 400-449 | (None) |
| 350-399 | Rehoboth, Newark Grammar, Dover |
| 300-349 | Shelbyville, Milford, Lewes, Milton, Georgetown, Smyrna, Middletown, New Castle Elementary, Seaford, Delaware City |
| 250-299 | Felton, Laurel, Delmar |
| 200-249 | Newark Academy, Millsboro, Frankford |
| 150-199 | Clayton, New Castle Academy, Dagsboro |

The rural schools, however, show even more deplorable conditions; and since school surveys have generally treated of city school buildings, it seems best to analyze the situation in detail for these rural communities. Only schools for white children will be considered. In Table II are given the scores assigned to each of the two hundred and ninety-three village and rural schools for such children—to the schools, in other words, having one to four teachers.

TABLE II. SCORES FOR ONE- TO FOUR-TEACHER SCHOOLS
FOR WHITE CHILDREN
(Maximum Possible Score, 1,000)

| Scores of Buildings | No. of Buildings |
|------------------------|---------------------|
| 0- 49 | 0 |
| 50- 99 | 1 |
| 100-149 | 6 |
| 150-199 | 21 |
| 200-249 | 72 |
| 250-299 | 90 |
| 300-349 | 49 |
| 350-399 | 31 |
| 400-449 | 15 |
| 450-499 | 5 |
| 500-549 | 2 |
| 550-599 | 0 |
| 600-649 | 0 |
| 650-699 | 1 |
| Total..... | 293 |
| 25-percentile..... | 230.77 |
| Median..... | 275.00 |
| 75-percentile..... | 329.35 |

This table reads as follows: one building scored between 50 and 99 points, and six buildings scored between 100 and 149 points. As is indicated at the bottom of the column headed "Number of Buildings," one quarter of the whole number of buildings scored less than 231 points, one-half of them less than 275 points, and three-quarters of them less than 330 points.

An analysis of the scores given to these schools with respect to site, building, service systems, classrooms, and special rooms will make clear the need for the reconstruction of nearly all of these buildings on a new site.

With reference to site, it will be observed that, according to Table III, one-half of the buildings scored less than 66 points out of a possible 160. Stated in terms of the actual situation, it may be said that the typical school site in Delaware is a piece of

ground, triangular or rectangular in form, with an area of less than one-half an acre, in the angle of a crossroads. It is ordinarily not considered suitable for cultivation; and it is not infrequently so low as to make it difficult or impossible to drain it properly. In a great many cases this little plot of ground is either in a densely wooded section or on the edge of a forest or swamp. Village and rural schools in Delaware need larger and better school sites. Playgrounds, play apparatus, and teachers who can teach children how to play, are needed throughout the state. The survey has revealed the need, and the people will, it is hoped, provide the sites.

TABLE III. SCORES FOR ONE- TO FOUR-TEACHER
SCHOOLS FOR WHITE CHILDREN WITH
RESPECT TO SCHOOL SITE
(Maximum Possible Score, 160)

| Scores | No. of Buildings |
|----------------------|---------------------|
| 0- 9 | 0 |
| 10- 19 | 2 |
| 20- 29 | 7 |
| 30- 39 | 20 |
| 40- 49 | 19 |
| 50- 59 | 38 |
| 60- 69 | 63 |
| 70- 79 | 37 |
| 80- 89 | 40 |
| 90- 99 | 31 |
| 100-109 | 16 |
| 110-119 | 14 |
| 120-129 | 3 |
| 130-139 | 3 |
| Total. | 293 |
| 25-percentile. . . . | 55.32 |
| Median. | 66.00 |
| 75-percentile. . . . | 86.75 |

A maximum possible score of 200 points is allowed for the building, as such, including its placement on the site, its gross structure, and its internal structure. Gross structure includes the type of building, material used, height, roof, foundation, walls, entrances, and the like. Internal structure includes stairways, vestibules, corridors, basement, color scheme, and the like. The scores for these elements are indicated in Table IV.

TABLE IV. SCORES FOR ONE- TO FOUR-TEACHER SCHOOLS
FOR WHITE CHILDREN WITH RESPECT TO
GROSS STRUCTURE OF BUILDING
(Maximum Possible Score, 200)

| Scores | No. of Buildings |
|-----------------------|---------------------|
| 0- 9 | 0 |
| 10- 19 | 3 |
| 20- 29 | 4 |
| 30- 39 | 21 |
| 40- 49 | 38 |
| 50- 59 | 59 |
| 60- 69 | 60 |
| 70- 79 | 48 |
| 80- 89 | 19 |
| 90- 99 | 14 |
| 100-109 | 12 |
| 110-119 | 6 |
| 120-129 | 5 |
| 130-139 | 2 |
| 140-149 | 1 |
| 150-159 | 0 |
| 160-169 | 1 |
| Total | 293 |
| 25-percentile | 50.23 |
| Median | 62.66 |
| 75-percentile | 76.09 |

It will be observed by reading this table that one-half of the buildings scored less than 63 points out of a total of 200. Most

of the buildings in Delaware are very old frame, stone, or brick structures, that have outlived their usefulness. The typical building is a most unattractive, store-box type of structure, with windows located at regular intervals on two or three sides, with the entrance on the fourth side. In almost no case is there any evidence of an attempt to make the school building attractive architecturally. In many of these buildings there is not even a vestibule. In more than half of them no cloakroom is provided. In many instances the roof, sadly in need of repair, has permitted enough water to enter the building to damage the plaster, even to the point of causing a large part of it to fall off the ceiling and walls. In some cases a leaking roof has resulted in a moldy interior, totally unfit for school purposes. Floors are badly damaged, books are moldy, and school furniture is discolored and covered with dirt on account of a failure to provide adequate protection against the rain.

The most striking failure to provide adequately for these school children is revealed in the scores which were allowed for service systems. When one remembers that this includes heating and ventilation, fire protection, cleaning, artificial lighting, water supply system, and toilets, it is easy to understand what a very low score means.

From Table V it will be observed that one-half of the school buildings received a score of less than 55 points out of a possible 250. More than half of these buildings had no source of heat other than the old-fashioned stove placed in the middle of the room. Only three buildings were provided with inside toilets. The toilets placed on the school grounds were almost uniformly in a most unsanitary condition, and were very generally marked with obscene pictures and writing. If the children of the state are to be properly cared for and freed from the contaminating influence of the outhouses now found everywhere on the grounds, toilet rooms on the same floor as the classrooms, with sewer connections, sewage disposal, or mechanical toilet tank, should be installed. In the buildings which are to be constructed a modern form of jacketed stove should be provided; or, if the consolidations proposed by the survey commission are made, heating plants in the basement should be installed.

TABLE V. SCORES FOR ONE- TO FOUR-TEACHER
SCHOOLS FOR WHITE CHILDREN WITH RESPECT
TO SERVICE SYSTEMS
(Maximum Possible Score, 250)

| Scores | No. of Buildings |
|--------------------|---------------------|
| 0- 9 | 0 |
| 10- 19 | 2 |
| 20- 29 | 17 |
| 30- 39 | 48 |
| 40- 49 | 52 |
| 50- 59 | 50 |
| 60- 69 | 71 |
| 70- 79 | 31 |
| 80- 89 | 13 |
| 90- 99 | 5 |
| 100-109 | 3 |
| 110-119 | 0 |
| 120-129 | 0 |
| 130-139 | 0 |
| 140-149 | 0 |
| 150-159 | 0 |
| 160-169 | 0 |
| 170-179 | 0 |
| 180-189 | 0 |
| 190-199 | 1 |
| Total..... | 293 |
| 25-percentile..... | 40.18 |
| Median..... | 54.50 |
| 75-percentile..... | 66.24 |

Space devoted to classroom use is allowed a score of 225 points out of a thousand. Under this head are included the size and form of the classroom, its floors, walls, ceiling, and their condition, the lighting of the room, the furniture, the cloakroom and wardrobes adjacent to or included in the classroom, blackboards, pictures, and teaching equipment.

TABLE VI. SCORES FOR ONE- TO FOUR-TEACHER
SCHOOLS FOR WHITE CHILDREN WITH RESPECT
TO CLASSROOMS
(Maximum Possible Score, 225)

| Scores | No. of Buildings |
|--------------------|---------------------|
| 0- 9 | 0 |
| 10- 19 | 0 |
| 20- 29 | 0 |
| 30- 39 | 1 |
| 40- 49 | 3 |
| 50- 59 | 13 |
| 60- 69 | 31 |
| 70- 79 | 40 |
| 80- 89 | 47 |
| 90- 99 | 69 |
| 100-109 | 30 |
| 110-119 | 15 |
| 120-129 | 17 |
| 130-139 | 11 |
| 140-149 | 5 |
| 150-159 | 5 |
| 160-169 | 1 |
| 170-179 | 2 |
| 180-189 | 1 |
| Total..... | 293 |
| 25-percentile..... | 74.75 |
| Median..... | 90.34 |
| 75-percentile..... | 103.82 |

It will be observed from reading Table VI that one-half of the schools were scored less than 90 points out of a possible 225. This is due to the fact that in most cases the classroom was not sufficient in size, that in many cases the floor, walls, and ceiling were in miserable condition, that the lighting of the room was very seldom satisfactory, that the furniture was often of a very old type and seldom of a sufficient number of sizes to accom-

modate the children of the different age groups, and that in general there was a most remarkable absence of pictures, charts, globes, or other teaching equipment.

In order to conduct a modern school, rooms other than classrooms must be provided. One hundred and sixty-five points are allowed on the score card for special rooms, under which head are classified playroom, community room, officials' consultation room, library, lunch room, space for industrial and household arts, and fuel room. It is the contention of the survey commission that no one-room school provides the space necessary for the education of children of the varying ages found in a rural school. The scores for special rooms appear in Table VII.

TABLE VII. SCORES FOR ONE- TO FOUR-TEACHER
SCHOOLS FOR WHITE CHILDREN WITH RESPECT
TO SPECIAL ROOMS
(Maximum Possible Score, 165)

| Scores | No. of Buildings |
|-----------------------|---------------------|
| 0 | 33 |
| 1 | 27 |
| 2 | 103 |
| 3 | 92 |
| 4 | 13 |
| 5 | 11 |
| 6 | 1 |
| 7 | 2 |
| 8 | 0 |
| 9 | 1 |
| 10-19 | 2 |
| 20-29 | 4 |
| 30-39 | 1 |
| 40-49 | 3 |
| Total | 293 |
| 25-percentile | 1.12 |
| Median | 1.84 |
| 75-percentile | 2.60 |

The reading of the table indicates that a very minimum of such equipment has been provided for the one- to four-teacher schools of Delaware. Only 14 buildings out of the 293 rural schools for white children received more than 5 points out of a possible 165.

Certain other interesting facts with respect to these buildings are indicated in Table VIII. From this table it will be observed that 45 percent of the buildings are located on a triangular or irregular site, that 30 percent of them are without vestibules, and so on for the other items.

TABLE VIII. NUMBER AND PERCENT OF BUILDINGS
WITH SPECIFIED DEFECTS AND FACILITIES

| | No. of Buildings | Percent |
|--|---------------------|---------|
| Located on triangular or irregular site..... | 131 | 45.0 |
| Without vestibule..... | 87 | 30.0 |
| No water supply..... | 81 | 28.0 |
| Without improved type of heater..... | 134 | 45.7 |
| Without flagpole..... | 120 | 41.0 |
| Without cloakrooms..... | 183 | 62.5 |
| Low in glass area..... | 144 | 49.1 |
| Light from 1 side..... | 0 | 0 |
| " " 2 sides..... | 165 | 56.3 |
| " " 3 sides..... | 102 | 35.0 |
| " " 4 sides..... | 15 | 5.1 |
| " " 5 sides..... | 0 | 0 |
| " " 6 sides..... | 1 | 0.3 |
| " " 7 sides..... | 1 | 0.3 |
| With gas or electricity..... | 4 | 1.4 |
| With indoor toilets..... | 3 | 1.0 |
| With playroom..... | 4 | 1.4 |
| With community room..... | 5 | 1.7 |
| With office..... | 1 | 0.3 |
| With industrial arts..... | 0 | 0 |
| With household arts..... | 0 | 0 |
| With library..... | 1 | 0.3 |
| With lunch room..... | 1 | 0.3 |

A similar set of tables of scores for the buildings used by the colored children of the state would indicate an even more deplorable situation than that revealed in the analysis of the provisions made for the white children.

The hopeful part of the entire situation for Delaware, however, is to be found in the generosity of Mr. du Pont and in the action already taken by a great many communities to abandon the buildings which are unfit for use and to construct modern schools buildings on adequate sites. These new buildings will cost more than has heretofore been expended on school buildings in the state of Delaware. If the recommendations of the survey commission are followed, the number of one-teacher schools will be diminished very rapidly on account of the consolidations which will be effected. There may remain a very few one-teacher schools, but even in these remote communities adequate facilities should be provided, even though the cost be greater; and unusually capable teachers should be provided for them, even though it may be necessary to pay a larger salary than is ordinarily paid to those who teach in the towns and cities of the state.

A similar survey of the school buildings of other states might show either a better or a worse condition. It can be confidently stated, however, that any such survey would make clear the necessity for the development of a state-wide school-building program such as has been undertaken in no state except Delaware.

SHORTER METHOD FOR COMPUTING THE COEFFICIENT OF CORRELATION

LEONARD P. AYRES

Russell Sage Foundation, New York City

The coefficient of correlation measures the agreement between two series of numbers. The numbers are in pairs and each number in the first series is the first member of a pair of measurements, while the corresponding number in the second series is the second member of that pair. The object of computing the coefficient of correlation is to find in what degree smaller quantities in one series imply correspondingly smaller quantities in the other, and larger quantities in the one imply correspondingly larger quantities in the other. The coefficient of correlation does this by comparing the amount of agreement that does exist with the perfect agreement that might exist.

THE PEARSON COEFFICIENT

The formula for finding the Pearson coefficient of correlation is best expressed as:

$$r = \frac{\Sigma (x.y)}{\sqrt{\Sigma x^2} \sqrt{\Sigma y^2}}$$

In this formula:

r = correlation

Σ = summation

x = deviations from the average of the first series

y = deviations from the average of the second series

The use of the formula may be illustrated by finding the correlation between two short and simple series as follows:

| | |
|----|----|
| 3 | 6 |
| 4 | 5 |
| 5 | 7 |
| 8 | 11 |
| 10 | 11 |
| — | — |
| 30 | 40 |

Since there are five numbers in each series, the average of the first series is 6, and that of the second is 8. The next step, according

to the formula, is the finding of the deviations of these numbers away from their averages. Calling these deviations x and y they are as follows:

| x | y |
|-----|-----|
| -3 | -2 |
| -2 | -3 |
| -1 | -1 |
| +2 | +3 |
| +4 | +3 |

It is clear that there is considerable agreement between the general trends of these two series. Large minus deviations in the first series are accompanied by large minus deviations in the second series and large plus deviations in the first by large plus deviations in the second. The problem is to find how nearly complete this agreement is; how near the partial agreement that does exist comes to equaling the perfect agreement that might exist.

THE ACTUAL DEVIATIONS

This involves comparing the actual deviations of these two series with those that might be found if the agreement were perfect. The actual deviations are measured by multiplying the corresponding members of the two series and getting the sum of these products as follows:

| xy |
|-------|
| +6 |
| +6 |
| +1 |
| +6 |
| +12 |
| <hr/> |
| 31 |

DEVIATIONS FOR PERFECT AGREEMENT

The next step is to find a measure of the deviations that would exist if the agreement between the two series were perfect. The first deviation of the x series is -3 and the second is -2. If the corresponding deviations of the second series were identical they would also be -3 and -2 and there would be perfect agreement.

Following the method already used in measuring the actual deviations, a measurement may be found for the deviations that would exist if all those of the second series were equal to those

in the first. This is done by multiplying all these deviations of the first series by themselves or finding their squares. The same process is now carried through with the second series and the squares of both series of deviations are the following:

| x^2 | y^2 |
|-------|-------|
| 9 | 4 |
| 4 | 9 |
| 1 | 1 |
| 4 | 9 |
| 16 | 9 |
| — | — |
| 34 | 32 |

We now have three quantities: 31 which measures the actual deviations of the two series, 34 which shows what the deviations of the x series would amount to if they were accompanied by identical deviations in the y series, and 32 which shows what those of the y series would amount to if they were accompanied by identical deviations in the x series.

COMPARING THE ACTUAL WITH THE POSSIBLE

The actual deviations must now be compared with those that might exist if the correlation were perfect. This is done by comparing 31 with 34 and 32 simultaneously and, since 34 and 32 are the sums of series of squares, the comparison is made by finding the product of their square roots and comparing 31 with that number. This working out of the formula gives the following:

$$r = \frac{\Sigma x \cdot y}{\sqrt{\Sigma x^2} \sqrt{\Sigma y^2}} = \frac{31}{\sqrt{34} \sqrt{32}} = \frac{31}{32.98} = 0.94$$

The work described is easy in the artificially simple example chosen but far more laborious in cases where the average is not a whole number, for this involves multiplying and squaring with numbers having decimal fractions. There are methods for doing this that involve finding the deviations from a whole number near to the average and then correcting the work to allow for the differences between the results so found and the correct results. These methods are laborious and increase the chances of letting errors of computation destroy the accuracy of the work.

A SHORTER METHOD

All these difficulties may be avoided by a method which it is the object of this article to present, and which gives the sums of

the products and the sums of the squares of the deviations directly from the squares of the original numbers. It does not involve arranging the series in order, or finding the separate deviations, and so avoids the necessity for taking into account the plus and minus signs of deviations.

This method is based on considering every number in a series as being equal to the average of the series plus a plus or minus deviation from that average. Thus 10, which is the last number in the first series of the illustrative example already cited, may be considered as being equal to the average of the series, or 6, plus the deviation 4.

If now we find the sum of the squares of the numbers in a series and subtract from it the product found by multiplying the square of the average of the series by the number of cases, the remainder will be the sum of the squares of the deviations from the average. Suppose, for example, that the series consists of the two numbers 4 and 8. Their average is 6 and the deviations from it are -2 and +2, the sum of the squares of which equals 8. But it is also true that the squares of these numbers are 16 and 64 with a sum of 80. If we subtract from this the square of the average, or 36, multiplied by the number of cases, or 2, we again get 8 as a result.

THE FORMULA

Utilizing this principle we may compute coefficients of correlation directly from the products and squares of the items of the two series without finding the separate deviations. Before writing the formula certain abbreviations may be set down as follows:

- Sub. = the first, or subject series
- Rel. = the second, or relative series
- Tot. = the total of a series, the sum of its items
- Av. = the average of a series

The formula expressed in these abbreviations, and also using the letters *S* and *R* for subject and relative, is as follows:

$$r = \frac{\text{Sum of products of subject and relative items} - \text{av. of subject items} \times \text{tot. of relative items}}{\sqrt{(\text{Sum of squares of sub. items} - \text{av.} \times \text{total of these items}) (\text{Sum of squares of rel. items} - \text{av.} \times \text{total of these items})}}$$

It may be more easily written but perhaps less easily remembered if expressed as follows:

$$r = \frac{\Sigma S \cdot R - \frac{\Sigma S \cdot \Sigma R}{N}}{\sqrt{\left[\Sigma S^2 - \frac{(\Sigma S)^2}{N} \right] \left[\Sigma R^2 - \frac{(\Sigma R)^2}{N} \right]}}$$

The application of this formula to the example already worked by the conventional method is as follows:

| | Sub. | Rel. | Sub. ² | Rel. ² | Sub. × Rel. |
|---------|------|------|-------------------|-------------------|-------------|
| | 3 | 6 | 9 | 36 | 18 |
| | 4 | 5 | 16 | 25 | 20 |
| | 5 | 7 | 25 | 49 | 35 |
| | 8 | 11 | 64 | 121 | 88 |
| | 10 | 11 | 100 | 121 | 110 |
| | — | — | — | — | — |
| Total | 30 | 40 | 214 | 352 | 271 |
| Average | 6 | 8 | | | |

First correction = $271 - (6 \times 40) = 271 - 240 = 31$

Second correction = $214 - (3 \times 60) = 214 - 180 = 34$

Third correction = $352 - (8 \times 40) = 352 - 320 = 32$

$$r = \frac{31}{\sqrt{34 \times 32}} = \frac{31}{32.98} = 0.94$$

An important feature of this method is that it facilitates the use of mechanical aids to computation. The work consists almost entirely of taking squares from tables and multiplying and adding, which are the operations most easily done by machines. It eliminates the long series of subtractions which are not readily done by the help of machines. It also eliminates the necessity for watching for unlike signs in the multiplication of deviations, and making the series of corrections to allow for the differences between the guessed averages and the true averages of the series.

Experience in using this formula in dealing with the problems arising in the regular work of a statistical office indicates that it reduces by something more than one-half the time necessary to compute correlations from short series of 20 or 30 pairs of items. On longer series the saving is greater and on distribution tables involving some thousands of cases, grouped in about 100 compart-

ments, it cuts the time to one-third or one-fourth of that required by the old methods. Almost more important is the fact that its use very greatly reduces the number of errors made in working out the computations.

The use of the theorem that the sum of the squares of the deviations of a series of items away from their average is equal to the sum of the squares of the items minus the product of the total of the series and its average will greatly facilitate the finding of the standard deviation. Those who are interested in seeking a further explanation of the theorem itself will find it in considering each item of the series as a binomial consisting of the average plus a plus or minus deviation from the average. By squaring a series of such binomials and adding them, it will be found that the middle terms cancel out, leaving the squares of the average and the squares of the deviations. A similar procedure in multiplying the items of the two paired series will explain the process followed in deriving the numerator of the fraction given in the new formula.

Editorials

TEXTBOOKS: THEIR COST AND IMPROVEMENT

It is generally agreed that the results, the methods, and the spirit of research ought to be evident in school practice. As long as they are confined to remote laboratories and described in technical treatises, they merely exist in a vacuum. We, therefore, seek to bring them to the attention of those who supervise and teach. But this is not sufficient. The practice of teachers is largely controlled by the textbook. As is frequently pointed out, it plays a unique part in the affairs of the American school. It is, therefore, highly important that the textbook should exemplify the best that competent research has established. There should be an easy transfer of the findings of investigators to the content of schoolbooks. It is not enough to train the teacher in the spirit of research; the textbook should be the concrete fruition of research.

This is not often true. Constant revision of existing texts and frequent publication of new ones can only be accomplished if boards of education set aside considerably more money for books than they now do. There must not only be a larger fund for teachers' salaries, but there must also be a larger textbook fund. In fact, the whole scheme of universal public education can no longer be effectively maintained without going to new and higher cost levels.

Expenditures for textbooks in this country are very small, considering their importance in the teaching process. They cost rather more than perfumery and cosmetics and rather less than chewing gum. In 1913, the aggregate of sales for elementary and high-school use was a little over \$14,000,000. This amounted to 78 cents for each child enrolled in the public schools, to 57 cents for each child on the census rolls, and to somewhat less than 15 cents per capita of the population. Doubtless, the figures are a little higher now, but we all know that due to long-time contracts,

the cost of textbooks has been held down. At any rate, these figures are of such an order of magnitude that no change which could possibly have taken place in the last six or seven years would make any material difference in our argument. Average conditions for 1914 in 21 states showed that if every pupil who entered the first grade remained through the eighth grade, a complete layout of new basic textbooks for his eight-year course would cost but \$9.40. It is said that we are an extravagant people, that we are passing through a period of wild expenditure; but here is a place where our frugal nature apparently asserts itself. In the midst of the riotous living in which everyone except school teachers is supposed to be indulging, we are spending 15 cents a piece—or may it be 25 cents?—for the instruments of instruction with which we equip our schools.

It is not generally known, however, that our expenditures for this purpose are so meager. Exaggerated ideas on this subject seem to persist. In Illinois, for example, in 1912 wide publicity was given to an estimate that in that state the textbook profit amounted to \$1,000,000 per year. As a matter of fact, the total annual volume of business in Illinois was at that time only about \$750,000. Again, in connection with the political handling of the textbook question in Kansas a few years ago, wild estimates of what might be "saved to the tax payers" were made—estimates which sometimes exceeded the value of all the books sold in that state.

As a matter of fact, the cost of textbooks is so small that the real problem is not to cheapen them but to improve them and to increase the number available for each child. To do this we shall need to spend much more. Perhaps instead of remaining on the chewing gum basis, we shall have to lay out as much as we spend for soda water. For publishers cannot afford to incorporate new and better features into their books because the price of textbooks has been "standardized." Boards of education are willing to pay only about 32 cents for a speller, 90 cents for a history, or \$1.00 for a geography. With the introduction of free textbooks and the utilization of the large purchasing power of state, municipal, and county governments, this pseudo-standardization has been brought about—a condition whereby a speller is a speller and a history a history no matter what their covers may contain.

Investigations looking toward the improvement of textbooks are under way in a number of places and it is certain that this type of investigation will become more prevalent when its importance is better understood. Titles of recent masters' and doctors' theses make clear the fact that in departments of education in colleges and universities, the two fields of research which are regarded as most significant are the study of teaching methods and the analysis of curriculums. In the public schools, however, it is equally clear that the teaching methods and curriculum actually put into practice are the methods and curriculum of the textbook. To be effective, therefore, it is in the textbooks that the new methods and curriculums must be expressed.

It has been pointed out that three financial obstacles stand in the way: an inverted sense of values, an exaggerated notion of textbook costs, and a wrong-headed idea of efficiency. The inverted sense of values is evident in our vast expenditures for trumpery while we each contribute the price of a movie ticket that twenty million pupils may have the benefit of the most effective instruments of instruction. The exaggerated notion of costs is locally present wherever textbooks instead of being provided at public expense are purchased by parents; and the notion has also become generally current because of political rodomontade and reports (true and false) of crooked adoptions. The wrong-headed idea of efficiency is due to lay purchasing agents whose aim is to make a financial record and who judge the value of a book not by its contents but by the class to which it belongs.

For these reasons, although we are in the midst of researches whose results ought to be at once reflected in better textbooks, better textbooks are slow to appear. If it is a day of promise in research, it is also a day of stagnation in textbook making. The results, the methods, and the spirit of research will never become fully applicable in the work of the schools until in addition to their apprehension by teachers, they find expression in the books which are placed in the hands of pupils.

B. R. B.

Reviews and Abstracts

STRAYER, GEORGE D., ENGELHARDT, N. L., AND HART, F. W. *General report on buildings and grounds of Delaware, Wilmington, Delaware: Service Citizens of Delaware, 1919.* 222 pp.

One or more of the authors of the report visited every school house in Delaware, rated it by means of the well-known Strayer score card for school buildings, made observations concerning those features of each building to which they thought the attention of the people concerned should be called, and submitted a complete report to the Service Citizens of Delaware. The printed report gives representative cases for each type of building. Concerning each case reported there appear, first, the comments concerning the special needs of the particular building; second, the scores given the building on (a) site, (b) building, (c) service systems, (d) classrooms, and (e) special rooms; and third, in most cases a photograph. Chapter I contains a rather full statement of the characteristics of a modern school building. Part II gives a series of definitions of so-called minimum standards for each aspect of a school building listed on the score card, and tables of distribution of Delaware's buildings by counties according to scores on each of the main headings.

To the general reader residing outside of Delaware, the most significant part of the report is Chapter I. In it is given a brief but complete description of the essentials of a modern school building. The language is simple and direct. This splendid summary of modern requirements in school buildings should prove of great value to school boards as well as to superintendents at any time; but it is particularly timely just now when so many communities of all types are inaugurating school building programs. These modern requirements are naturally far in advance of actual building conditions not only in Delaware but also, no doubt, in all other states. Nevertheless, some buildings in Delaware are found to measure up well with these requirements, and there is a delightful assumption on the part of the authors that all the districts of Delaware will naturally wish to reach these high standards when they are made aware of them.

To the student of education there is much food for thought in Part II. The attempt to define minimum standards with respect to each subhead appearing on the score card brings out both the utility and the difficulty of such an undertaking. The authors have met the problem of difference in judgment which is bound to arise in respect to many of the standards by adopting the term "Optional" for many of the most debatable essentials. For example, under "Seats and Desks," they give as minimum, "Should be individual, adjustable, and *adjusted*," but they add as optional, "Movable

chairs are preferable to screwed-down seats." In spite of the inevitable differences of opinion concerning the definitions of these minimum standards, the well-classified and carefully worded definitions will prove very helpful indeed as a practical guide.

The genuine value of a score card as a method of evaluating buildings is clearly brought out in the distribution tables. Here we see a figure standing for the total score of a building, just as a figure stands for the score of a thoroughbred animal. One building can be compared with another, either as a building or in respect to any of its aspects. Comparison can also be made between the whole set of Delaware buildings and any other group of buildings measured by the same score card. Probably no other method of stating building facts could have been as effective in bringing the state to a realization of the seriousness of its building situation.

The chapters dealing with the individual buildings should prove of great value to the districts concerned. Even though the situation revealed is often deplorable, the account is only rarely couched in terms of rebuke. The several districts cannot help feeling that the survey was made for a constructive purpose, and will undoubtedly utilize many of the suggestions in constructing new buildings.

F. J. K.

OBRIEN, FRANCIS P. *The high school failures.* (Teachers College Contributions to Education, No. 102.) New York: Teachers College, Columbia University, 1919. 97 pp.

The free public high school enrolls about 85 percent of all the secondary school pupils in the United States. That fact alone makes appropriate a study of the products in terms of success or failure in attaining the standards set by particular schools. This book is such a study. In it the author analyzes the cumulative records of 6,141 pupils in eight different high schools in New York and New Jersey, in order to determine the significant facts relative to pupils who fail. For this group of pupils he has answered definitely and conclusively these questions which are stated as chapter headings:

"How extensive are the failures of the high school pupils?"

"What basis is discoverable for a prognosis of the occurrence or the number of failures?"

"How much is graduation or the persistence in school conditioned by the occurrence or by the number of failures?"

"Are the school agencies employed in remedying the failures adequate for the purpose?"

"Do the failures represent a lack of capability or of fitness for high school work on the part of those pupils?"

"What treatment is suggested by the diagnosis of the facts of failure?"

Some of the more important findings are:

1. The percent of recorded and virtual but unrecorded failures is 66.
2. Of the graduating pupils 58.1 percent fail one or more times.
3. The number of failures increases for the first four semesters, with one-third to one-half of the pupils failing in each semester to the seventh.

4. Of all the failures 86 percent occur for pupils from 15 to 18 years of age, inclusive.

5. Mathematics, Latin, and English combined furnish 60 percent of the total failures, but English furnishes many rather because of large enrollment than because of difficulty. The percent of failures of pupils taking the subject is greatest for mathematics, Latin, and German.

6. The younger entering pupils have fewer failures and are more certain of graduation.

7. The first year's record has real prognostic value for pupils who remain in school more than three semesters.

8. As many graduates as non-graduates fail—31.5 percent in each case.

9. "Failure is probably not a prime cause of dropping out for most of the non-graduates, as 80 percent have only five failures or fewer." (p. 58)

10. "The worst consequences of failure are perhaps in acquiring the habit of failing and in coming to accept oneself as a failure." (p. 58)

11. "The emphasis placed on repetition is excessive and the faith displayed in it by school practice is unwarranted by the facts." (p. 74)

12. The number of pupils who lack native ability is small.

Thus the author does not regard lack of intellectual ability as being a particularly potent cause of failure. He rightly points out that the pupil is only one of the factors contributing to failure. His emphasis in the matter of obviating failure is placed upon subject matter and the aims of teaching.

In a sense the author is right. From the very nature of the case, as Burk points out, there can be no misfit children. There may be misfit courses of study, misfit teachers, or a misfit educational philosophy, but there can be no misfit child because all these must be fitted to the child. As well say that a person does not fit his clothes as to say that a child does not fit the school.

On the other hand, one should point out that the trend of thought today is in the direction of emphasizing intelligence as an important cause of failure and of leaving school. Professor Terman and his students are bringing to bear a large amount of data on this question. They are doing so by actually administering intelligence tests to the children who fail as well as to those who do not. One may point out that Dr. Obrien's conclusion that failing pupils "lack neither ability nor earnestness" is not based on evidence as direct and convincing as is the evidence gathered from actual testing for intelligence.

As remedies for failure, Dr. Obrien suggests, among other things: a functioning psychology; whole-hearted purposeful responses from the pupil; properly selected student advisors; a high degree of flexibility in the curriculum; the reduction of subject matter to smaller units; provision for directed study; and the collection and interpretation of useful facts. Some of these suggestions are important, others are rather naive; and all are more or less without foundation in evidence.

As needed analyses of this kind are made and the unescapable facts of failure are produced, the important question for secondary education is to determine what is to be done about it. Is the high school to be a selective institution for the purpose of educating the leaders of society? Or is it to

be an institution where every youth who can profit from secondary instruction is enrolled? Just now we are at the parting of the ways with the traditional sign post pointing to a continuance of absolute uniformity. The new re-interpreted social theory demands that each individual shall live as fully as we know how to permit him. In practice, therefore, curriculums must be devised so as to find out the interests, ideals, abilities, and even the prejudices of pupils; and then provision must be made for the full development of interests, ideals, and abilities.

P. E. BELTING

PUBLICATIONS RECEIVED

- Lennes, N. J. and Jenkins, Frances, *Applied arithmetic: three essentials, books 1, 2, 3*. Philadelphia, J. B. Lippincott Co., 1919. 3v.
- Obrien, Francis P. *The high school failures*. (Teachers College Contributions to Education, No. 102.) New York: Teachers College, Columbia University, 1919. 97 pp.
- Payne, E. George. *Education in Accident Prevention*. Chicago: Lyons and Carnahan, 1919, 158 pp.
- Powell, Lyman P. and Hodgins, Fred B. *America and the league of nations*. Chicago: Rand McNally and Co., 1919. 208 pp.
- Powell, Lyman P. and Powell, Gertrude W. *The spirit of democracy*. Chicago: Rand McNally and Co., 1918, 272 pp.
- Powell, Lyman P. and Curry, Charles M. *The world and democracy*. Chicago: Rand McNally and Co., 1919. 553 pp.
- Pyle, Wm. Henry. *The science of human nature*. Boston: Silver, Burdette and Co., 1917. 229 pp.
- Rusmisl, L. C. *Industrial-commercial geography of the United States*. New York: The A. N. Palmer Co., 1918. 294 pp.
- Sanford, Chester M. and Owen, Grace A. *Modern Americans—a biographical school reader for the upper grades*. Chicago: Laurel Book Company, 1918. 203 pp.
- Serl, Emma and Pelo, Wm. J. *American ideals*. New York: The Gregg Publishing Company, 1919. 160 pp.
- Smith, David Eugene. *Number stories of long ago*. Boston: Ginn and Co., 1919. 136 pp.
- Smith, Reginald Heber. *Justice and the poor*. New York: The Carnegie Foundation for the Advancement of Teaching, Bulletin No. 13, 1919. 271 pp.
- Spillman, Harry Collins. *Personality. Studies in personal development*. New York: The Gregg Publishing Co., 1919. 206 pp.
- Tappan, Eva March. *An taryele-men history of our country*. Boston: Houghton Mifflin Co., 1914. 256 pp.
- Tappan, Eva March. *Our European ancestors*. Boston: Houghton Mifflin Co., 1918. 264 pp.
- Wilson, Guy M. *A survey of the social and business usage of arithmetic*. (Teachers College Contributions to Education, No. 100.) New York: Teachers College, Columbia University, 1919. 62 pp.

Bibliographies

A BIBLIOGRAPHY OF STANDARIZED TESTS FOR THE HIGH SCHOOL

WALTER S. MONROE

Assistant Director, Bureau of Educational Research, University of Illinois
(Continued from February Number)

IV. ENGLISH

A. Composition

Harvard-Newton Composition Scale. The Harvard-Newton Composition Scale consists of four separate scales, one for each form of discourse: argumentation, description, exposition, and narration. Each of the scales consist of six compositions written by eighth-grade pupils and arranged in order of merit as determined by the marks assigned by teachers, rating them as eighth-grade compositions. This scale has been used in the first-year classes in the high school.

Publisher: Harvard University Press, Cambridge, Massachusetts.

References: Ballou, F. W. *Scales for the measurement of English composition.* (Harvard Newton Bulletin No. 2) Cambridge, Massachusetts: Harvard University Press, 1914.

The derivation and use of the scale is described.

Stoddard, W. E. "Comparison of the Hillegas and Harvard-Newton scales in English composition," *Pedagogical Seminary*, 23:498-501, December 1916.

Second annual conference on educational measurements, April, 1915, p. 115 ff. (Indiana University Bulletin, v. 13, no. 11, October, 1915).

An account of the use of this scale in the first year of the high school is given.

Hillegas' Scale for the Measurement of the Quality in English Composition for Young People. This scale consists of ten compositions ranging from an artificial production, whose scale value is zero, to the tenth composition, whose scale value is 9.3. Three of the ten compositions are artificial productions, five were written by high school pupils, and the remaining two by college freshmen. No two were written on the same topic and they vary greatly in length and type. Each degree of merit is represented by only one composition.

Publisher: Bureau of Publications, Teachers College, Columbia University, New York City.

References: Curtis, S. A. *The Gary public schools: measurement of classroom products.* New York: General Education Board, 1919, p. 227.

An account of the use of this scale in the high school is given here and also in the following reference.

An educational survey of Janesville, Wisconsin. Madison, Wisconsin: State Department of Public Instruction, 1918, p. 287.

Gunther, Charles. "My experience with the Hillegas scale," *English Journal*, 2:535-42, November, 1919.

Describes a series of experiments carried on by four teachers in the use of the Hillegas scale. Mr. Gunther in his conclusion definitely states that through the Hillegas or some similar scale greater uniformity is gained than through a percent scale.

Hillegas, Milo B. "Hillegas scale for measurement of English composition," *Teachers College Record*, 13:331-84, September, 1912.

Johnson, F. W. "The Hillegas-Thorndike scale for measuring the quality in English composition by young people," *School Review*, 21:39-49, January, 1913.

Kayetz, Isidor. "A critical study of the Hillegas composition scale," *Pedagogical Seminary*, 21:559-77, December, 1914.

The first part of the article is concerned with the scale itself; the second part devoted to a criticism by the author and others. References relative to the scale follow.

Parker, F. E. "Value of measurements," *English Journal*, 8:203-17, April, 1919.

A criticism on the use of the Hillegas scale by a subcommittee of the Detroit English Club. S. A. Courtis discusses the "Use of the Hillegas Scale" in the second part of this article.

School report of the town of Southington, Connecticut, 1917, p. 12.

Stoddard, W. E. "Comparison of the Hillegas and Harvard Newton scales in English composition," *Pedagogical Seminary*, 23:498-501, December, 1916.

Most valuable for its criticism of the Hillegas scale.

Thorndike, E. L. "Notes on the significance and use of the Hillegas Scale for measuring the quality of English composition," *English Journal*, 2:551-561, November, 1913.

Minnesota English Composition Scale.

Reference: Van Wagenen, M. J. "The Minnesota English composition scale, its derivation and validity." To be published in the first volume of the *Journal of Educational Research*.

Nassau County Supplement to the Hillegas Scale. This scale is a simplification and improvement of the original Hillegas Composition Scale. It consists of a series of compositions written under controlled conditions and arranged in order of general merit.

Publisher: Bureau of Publications, Teachers College, Columbia University, New York City. Also obtainable from Bureau of Educational Research, University of Illinois, Urbana, Illinois.

References: Theisen, W. W. "Improving teachers' estimates of composition specimens with the aid of the Trabue Nassau County scale," *School and Society*, 7:143-150, February 2, 1918.

Discusses the outcomes of an experiment prompted by the question: "Can teachers improve their judgment of the merit of compositions through use of the scale?"

Trabue, M. R. "Supplementing the Hillegas composition scale," *Teachers College Record*, 18:51-84, January, 1917.

Reasons for constructing the Nassau County supplement, its advantages over the Hillegas scale. Table V lists the "estimates of quality to be expected from high school seniors and sophomores" in the use of the Hillegas scale.

An account of the use of scale in the high school may be found in the following:
Report of a Survey of Public Education in Nassau County, New York. Albany, New York: University of the State of New York, 1918, p. 149 ff.

Report of a survey of the school system of St. Paul, Minnesota, 1917, p. 456.

Scores are given for St. Paul, Minnesota; Mobile and Mobile County, Alabama; Nassau County, and South River, New York; Lead, South Dakota.

Topeka Scale for Judging Compositions. This scale was devised by the English department of the high school in Topeka, Kansas.

Address: Carmie S. Wolfe, High School, Topeka, Kansas.

Reference: "The Topeka scale for judging compositions," *Bulletin of the Illinois Association of Teachers of English*, v. 9 no. 6. March, 1, 1917.

Scale reproduced and derivation described. For a copy of the bulletin address Professor H. G. Paul, University of Illinois, Urbana, Illinois.

Willing's Scale for Measuring Written Composition. This scale consists of eight compositions arranged in order of both "form" and "story value." All of the compositions were written by school children under controlled conditions. This scale may be used in grades IV to XII inclusive.

Publisher: Bureau of Educational Research, University of Illinois, Urbana, Illinois.

Reference: Willing, M. H. "Measurement of written composition in grades IV to VIII," *English Journal*, 7:193-202, March, 1918.

Considers the derivation of the test, also its giving and scoring. Although this scale was designed for use in the upper grades of the elementary school, it can be used profitably in the high school.

B. COPYING

Boston Test in Accurate Copying. Test was prepared by the committee on standards in English. It was given to first-year pupils.

Reference: *English: determining a standard in accurate copying.* (Bulletin No. 6 of the Department of Educational Investigation and Measurement, Boston Public Schools, 1916).

This bulletin consists of a report on the Test in Accurate Copying given to 4,494 first-year high school pupils in November, 1914. Test is reproduced.

Kansas City, Missouri, public schools, bulletin no. 1 of the Bureau of Research and Efficiency, February, 1916, p. 55.

Analysis of errors made in Kansas City, Missouri, by ages and grades. The test also with directions for giving and scoring are given.

C. GRAMMAR AND PUNCTUATION

Clemens' Grammar Test. A test consisting of thirteen sentences in which the pupil is asked to make a choice between the correct and incorrect words. This test is based on E. C. Woolley's *Handbook of Composition*.

Address: A. H. Clemens, Principal of High School, Rochester, Minnesota.

Reference: *Report of the board of education of Rochester, Minnesota, 1912-1918*, p. 40.

The test is reproduced, with a key to the sections in Woolley's *Handbook*. Tentative standard scores are also given.

Starch's Grammatical Scales. Starch has devised three scales (A, B, and C) to measure a pupil's ability to use correctly certain language forms.

Starch's Tests for Measuring Grammatical Knowledge. The three tests in this series consist of selections in which the pupil is asked to indicate certain parts of speech or certain inflected forms.

Publisher: University Cooperative Company, Madison, Wisconsin.

References: Starch, D. *Educational measurements*, New York: Macmillan, 1916.

See Chapter VII entitled "The measurement of ability in English grammar," for a reproduction of the tests and standards.

Starch, Daniel. "The measurement of achievement in English grammar," *Journal of Educational Psychology*, 6:615-26, December, 1915.

Clemens' Punctuation Test. A list of twenty-eight sentences which are to be punctuated. Based on E. C. Woolley's *Handbook of Composition*.

Address: A. H. Clemens, Principal of High School, Rochester, Minnesota.

Reference: *Report of the board of education of Rochester, Minnesota, 1912-1918*, p. 42.

The test with a key to the sections in the *Handbook* is reproduced.

Starch's Punctuation Scale. This scale consists of a series of groups of sentences which are to be punctuated. The groups are arranged in order of difficulty. It will be useful in the high school.

Publisher: University Cooperative Company, Madison, Wisconsin. Also obtainable from the Bureau of Educational Research, University of Illinois, Urbana, Illinois.

References: Starch, Daniel. *Educational measurements*. New York: Macmillan Company, 1916.

See Chapter VII entitled "The measurement of ability in English grammar," for a reproduction of test, and standards.

Starch, Daniel. "The measurement of achievement in English grammar," *Journal of Educational Psychology*, 6:615-26, December, 1915.

D. LANGUAGE

Trabue's Completion-Test Language Scales. These scales are composed of "completion sentences" which are arranged in order of difficulty from the very simple to the very difficult. There are two scales for the high school, L and M.

Publisher: Bureau of Publications, Teachers College, Columbia University, New York City. Also obtainable from Bureau of Educational Research, University of Illinois, Urbana, Illinois.

References: Trabue, M. R. *Completion-test language scales*. (Teachers' College, Columbia University Contributions to Education, No. 77) New York: Teachers College, Columbia University, 1916.

The author states: "This monograph is the report of an attempt to derive one or more scales for the measurement of ability along certain lines closely related to language."

Account of its use in the high school may be found in:—

Report of a survey of public education in Nassau County, New York. Albany, New York: University of the State of New York, 1918, p. 177 ff.

Report of a survey of the school system of St. Paul, Minnesota, 1917, p. 433 ff., table inserted before p. 453, 460, 487.

E. SPELLING

Ayres' Spelling Scale. A thousand words constituting a fundamental English vocabulary have been grouped in this scale according to their spelling difficulty as indicated by the percent of correct spellings. The words were selected from four important reports on the frequency with which words are used in letters, newspaper articles, the English Bible, and various authors. The approximate percent of correct spellings to be expected in each grade is shown for each word. Although not intended for use in the high school this scale has been used by some above the eighth grade.

Publisher: Division of Education, Russell Sage Foundation, New York City.

Also obtainable from the Bureau of Educational Research, University of Illinois, Urbana, Illinois.

References: Ayres, L. P. *A measuring scale for ability in spelling*. New York: Division of Education, Russell Sage Foundation, 1915.

Courtis, S. A. *The Gary public schools: measurement of classroom products*. New York: General Education Board, 1919.

Scofield, F. A. "Difficulty of Ayres' spelling scale as shown by the spelling of 560 high school students," *School and Society*, 4:339-40, August 26, 1916.

Account of its use in the junior high school and senior high school, Eugene, Oregon.

Scofield, F. A. "An experiment in spelling in the Eugene high school," *School and Society*, 5:299-300, March 10, 1917.

Buckingham's Extension of the Ayres' Spelling Scale. This extension adds 509 words to the original Ayres' Spelling Scale making a scale of 1,509 words. The additions are mainly at the more difficult end of the scale and will thus add to its value especially for grammar grades and the high school. Scale is for use in grades II to IX.

Publisher: Bureau of Educational Research, University of Illinois, Urbana, Illinois.

Buckingham's Spelling Scale. This was the first published spelling scale. It consisted of fifty words very carefully evaluated together with about one hundred and twenty-five supplementary words. The scale as such has not been published separately. Some of the words now appear in the Buckingham's Extension of the Ayres' Spelling Scale mentioned above.

Publisher: Not obtainable.

Reference: Buckingham, B. R. *Spelling ability: its measurement and distribution.* (Teachers College, Columbia University Contributions to Education, No. 59). New York: Teachers College, Columbia University, 1913.

Considers the derivation of scale. Scale is reproduced.

An educational survey of Janesville, Wisconsin. Madison, Wisconsin: State Department of Public Instruction, 1918, p. 272.

Account of the Buckingham sixty-five word list in the high school.

Courtis' Dictation Spelling Tests. There are four tests in this series, each of which contains twenty test words embedded in sentences.

Publisher: S. A. Courtis, 82 Eliot Street, Detroit, Michigan.

Reference: Courtis, S. A. *The Gary public schools: measurement of classroom products.* New York: General Education Board, 1919, pp. 87-92.

Jones' One Hundred Spelling Demons. This is a list of the one hundred words most frequently misspelled by children in the elementary school.

Address: W. F. Jones, University of South Dakota, Vermillion, South Dakota.

Reference: Jones, W. F. *Concrete investigation of the material of English spelling* Vermillion, South Dakota: University of South Dakota, 1914.

This is an account of the investigation in which list was secured.

Monroe, Walter S., De Voss, J. C., and Kelly, F. J. *Educational tests and measurements.* Boston: Houghton Mifflin Company, 1917, p. 133.

The One Hundred Spelling Demons are reproduced.

Account of the use of the "Demons" in high school may be found in the following:

Annual report of the Des Moines public schools, 1915, part 2, p. 10 ff.

Annual report of the public schools of Johnstown, Pennsylvania, 1918, p. 111-113.

Courtis, S. A. *The Gary public schools: measurement of classroom products.* New York: General Education Board, 1919, p. 92 ff., 414 ff.

Jones' Spelling Demons used in analyzing misspelled words in the composition of eighth-grade children.

School report of the town of Southington, Connecticut, 1919, p. 21.

Annual report of the school department of the city of Pittsfield, Massachusetts, 1916, p. 23-24.

Monroe's Timed Sentence Spelling Test. This is a series of three tests each of which consists of fifty words taken from suitable columns of the Ayres' Spelling Scale and embedded in sentences. The sentences are then so arranged that they may be dictated at approximately the normal rate of writing in each of the grades. Test III is used in grades vii and viii and in the high school.

Publishers: Bureau of Educational Research, University of Illinois, Urbana, Illinois.

Monroe, Walter S., De Voss, J. C., and Kelly, F. J. *Educational tests and measurements.* Boston: Houghton Mifflin Co., 1917, p. 119.

V. FOREIGN LANGUAGES

Handschin's Foreign Language Tests. Series of four short tests for the purpose of discovering types of learners in language study.

Address: Charles H. Handschin, Oxford, Ohio.

Reference: Handschin, C. H. "A test for discovering types of learners in language study," *Modern Language Journal*, 3:1-4, October, 1918.

Tests reproduced and directions given.

A. French

Henmon's French Tests. A series of tests in French similar to the author's Latin tests.

Publisher: V. A. C. Henmon, University of Wisconsin, Madison, Wisconsin.

Reference: Henmon, V. A. C. "The measurement of ability in French," *Journal of Educational Research*. (To appear in the first volume.)

Starch's French Reading Test. This test is composed of a series of thirty sentences arranged roughly in the order of increasing difficulty. The sentences were from the readings usually covered in the first year of the high school.

Starch's French Vocabulary Test. The test is composed of two numbered lists of one hundred French words each and corresponding lists of their English equivalents. The pupil is asked to write after each French word a number of its English equivalents.

Publisher: University Cooperative Company, Madison, Wisconsin.

Reference: Starch, Daniel. *Educational measurements*. New York: Macmillan, 1916.

In Chapter XIII "The Measurement of Ability in French" the French vocabulary and sentence tests are reproduced. Derivation is described.

Wilkins' Predetermination Test in French. After analyzing language ability Mr. Wilkins derives a test for French which consists of five parts:

- (1) seeing and writing
- (2) hearing and writing
- (3) seeing and speaking
- (4) hearing and speaking
- (5) grammatical concepts.

Reference: Wilkins, L. A. "Testing for ability to learn a foreign language," *Bulletin of High Points in the Work of the High Schools of New York City*, v. 1, no. 2, p. 5, February, 1919; no. 8, p. 26, October, 1919.

A discussion of the test and scoring is followed by the reproduction of the test. Results are given in the October Number.

B. German

Starch's German Reading Test. This test, like the one for French, is composed of a series of sentences arranged roughly in the order of increasing difficulty. The sentences were selected from first-year texts and from the authors usually covered in four years of German.

Starch's German Vocabulary Tests. Professor Starch states: "The test is composed of two sets of 100 words each. These words are selected by taking the first word on every 23rd page of the large Muret-Sanders German-English Dictionary."

Publisher: University Cooperative Co., Madison, Wisconsin.

Reference: Starch, Daniel. *Educational measurements*. New York: Macmillan, 1916.

In Chapter XII: "The measurement of ability in German" tests are reproduced.

C. Latin

Brown's Latin Tests.

1. *Connected Latin Test.* This test consists of a connected passage of Latin. The pupils are given a specified amount of time in which to interpret and write in English as much of the passage as possible.
2. *Latin Sentence Test.* This consists of a series of Latin sentences ranging from very easy to very difficult.
3. *Formal Latin Vocabulary Test.* A list of words occurring in the Latin Sentence Test.
4. *Formal Latin Grammar Test.* This test is made up of twenty constructions in Latin sentences. The constructions are in italics and the pupils are required to name and describe them but not to translate the sentences.
5. *Functional Latin Grammar Test.* A series of Latin constructions chosen from the Latin Sentence Test. The pupils are graded on their ability to read correctly these constructions in their normal settings.

Publisher: Parker Educational Company, Madison, Wisconsin.

Reference: Brown, H. A. *A study of ability in Latin*. Oshkosh, Wisconsin: H. A. Brown, 1920.

Hanus' Latin Tests. These consist of four tests for vocabulary, a translation test and a grammar test. All of these tests are based on Caesar and Cicero. No words appear in the vocabulary tests "which occur less than one hundred times in Caesar and Cicero." The translation test "contains only constructions which are found at least 500 times in Caesar and Cicero."

Address: Paul Hanus, Harvard University, Cambridge, Massachusetts.

Reference: Hanus, P. H. "Measuring progress in learning Latin," *School Review*, 24:342-345, May, 1916.

Four Latin vocabulary tests are reproduced with suggested credit, directions for grading and making correlations.

Henmon's Latin Tests.

1. *Latin Vocabulary Tests* selected from 239 words common to thirteen beginners' books and to Caesar and Vergil.
 - (1) Texts A, B, C, and D are of equal difficulty and each consists of fifty words arranged in order of increasing difficulty.
 - (2) Text X is a selection of 5 words having approximately the same scale values, which may, therefore, be disregarded in scoring and each word given the same weight.
2. *Latin Sentence Tests.* Graded sentences containing no word not included in the 239 standard words.

(1) Tests I and II are of equal difficulty and each consists of ten sentences arranged in order of increasing difficulty.

(2) Test III is a selection of ten sentences of approximately the same difficulty.

Publisher: V. A. C. Henmon, University of Wisconsin, Madison, Wisconsin. Also obtainable from the Bureau of Educational Research, University of Illinois, Urbana, Illinois.

Reference: Henmon, V. A. C. "Measurement of ability in Latin," *Journal of Educational Psychology*, 8:515-38, November-December, 1917.

This article gives the derivation of the tests.

Lohr's Latin Test. This is a "form test" to be given at the end of the first year. Pupils are asked to identify the forms of certain nouns, pronouns, and verbs.

Address: School of Education, University of North Carolina, Chapel Hill, North Carolina.

Reference: Lohr, L. L. "A Latin form test for use in high school classes," *High School Journal*, 1:7-9, 14-17, November-December, 1918.

Gives critical account of the appropriateness of Starch's Latin tests and the derivation of a form test.

Starch's Latin Vocabulary Tests.

Starch's Latin Reading Test.

See Starch's French Tests for a description of these two tests.

Publisher: These scales have been withdrawn from distribution.

Reference: Starch, Daniel. *Educational measurements*. New York: Macmillan, 1916.

See Chapter XI: "Measurement of ability in Latin," for reproduction of tests, and their derivation.

Starch-Watters' Latin Tests. This is a second series of Latin tests prepared by Professor Starch, in collaboration with Mr. Watters. It consists of (1) Latin vocabulary test based on Lodge's Vocabulary of High School Latin; (2) translation tests for each of the four years of the high school course.

Address: University Cooperative Co., Madison, Wisconsin.

Wentworth Latin Tests. Test 1 consists of one hundred English words, some of which come from the Latin. The pupil is to mark each word showing whether it comes from the Latin or not. Test 2 consists of one hundred English words which are derived from the Latin. The pupil is to give the Latin root.

Address: Mary M. Wentworth, Hollywood High School, Los Angeles, California.

Reference: Wentworth, M. M. "An experiment with two Latin tests," *School Document No. 26*, Los Angeles City School District, Division of Educational Research, October, 1919.

D. Spanish

Wilkins' Predetermination Test in Spanish. Mr. Wilkins states that language ability is predicted largely upon (1) susceptibility to impression; (2) readiness and

accuracy of expression; (3) grasps of ordinary grammatical concepts. There are four ways for testing impression and expression:

1. Visual—motor (see and writing)
2. Aural—motor (hearing and writing)
3. Visual—oral (seeing and speaking)
4. Aural—oral (hearing and speaking).

Then there is a test of grammatical concepts.

Reference: Wilkins, L. A. "Testing for ability to learn a foreign language."

Bulletin of High Points in the Work of the High Schools of New York City, v. 1, no. 2, p. 5, February, 1919; October, 1919, no. 8, p. 26.

A discussion of the test and scoring is followed by the reproduction of the test. Results are given in the October Number.

VI. HANDWRITING

Ayres' Scale for Measuring the Quality of Handwriting of Adults. This scale is similar to the well-known Ayres' Scale, "three slant edition" except samples of the handwriting of adults were used in its construction.

Publisher: Russell Sage Foundation, Division of Education, New York City.

References: Ayres, L. P. *A scale for measuring the quality of handwriting of adults.*

(Division of Education, Russell Sage Foundation, Bulletin E 138.)

Penmanship: determining the achievement of elementary school graduates in handwriting (Bulletin No. IX of the Department of Educational Investigation and Measurement, Boston Public Schools, 1916.)

"On November 23, 1914, a test in accurate copying was given to 4,494 pupils in the first-year classes of fourteen high schools in Boston. In the test nothing was said to the pupils about their penmanship, hence they did not know that the quality of their handwriting was to be considered." The Ayres' Scale for Measuring Handwriting of Adults was used in rating these samples. Results are given.

Johnston, J. H. "A comparison of the Ayres and Thorndike handwriting scales (containing a table of equivalent values in the two scales)," *North Carolina High School Bulletin*, 7:170-73, October, 1916.

Ayres' Handwriting Scale, "Gettysburg Edition." This is the latest edition of Ayres' handwriting scales. Numerous changes have been introduced which make it a more accurate instrument than the "three slant edition." It uses the first three sentences of the "Gettysburg Oration" as copy material.

Publisher: Russell Sage Foundation, Division of Education, New York City.

Also obtainable from the Bureau of Educational Research, University of Illinois, Urbana, Illinois.

References: Breed, F. S. "The comparative accuracy of the Ayres' Handwriting Scale, Gettysburg Edition," *Elementary School Journal*, 18:459-63, February, 1918.

The "Gettysburg Edition" is shown to yield more accurate measures than the "three slant edition."

Lewis, E. E. "The present standard of handwriting in Iowa Normal Training High Schools," *Educational Administration and Supervision*, 1:663-71, December, 1915.

"This report describes an investigation of the quality and speed of the handwriting of 1,760 third and fourth-year students in 166 Iowa Normal Training High Schools."

Ayres' Measuring Scale for Handwriting. "Three Slant Edition." This is the first of the handwriting scales devised by L. P. Ayres.

Publisher: Russell Sage Foundation, Division of Education, New York City.

References: Ayres, L. P. *A scale for measuring the quality of the handwriting of school children.* (Division of Education, Russell Sage Foundation, Bulletin No. 113.)

This is an account of the derivation of the scale.

Courtis, S. A. *The Gary public schools: measurement of classroom products.* New York: General Education Board, 1919, p. 68.

Cleveland Free Choice Test. This refers only to the method of securing samples of the handwriting of pupils.

Reference: Courtis, S. A. *The Gary public schools: measurement of classroom products.* New York: General Education Board, 1919, p. 49.

VII. HISTORY

A. General

Rugg's Tests for Historical Judgment.

Address: E. U. Rugg, Oak Park High School, Oak Park, Illinois.

Reference: Rugg, E. U. "Character and value of standardized tests in history," *School Review*, 27:757-71, December, 1919.

The author discusses the existing tests in the field, general features and general criticisms of existing tests, value of the testing movement in history. He inserts a table which gives the following information: name and title, where reported, type, kind of questions and general comment.

B. Ancient History

Barnard's Test in Roman History. A test covering one period of ancient history.

Address: A. F. Barnard, University High School, University of Chicago, Chicago, Illinois.

Sackett's Scale in Ancient History. This scale is composed of six tests based on carefully selected information in ancient history.

Publisher: L. W. Sackett, University of Texas, Austin, Texas. Also obtainable from Bureau of Educational Research, University of Illinois, Urbana, Illinois.

Reference: Sackett, L. W. "A scale in ancient history," *Journal of Educational Psychology*, 8:284-93, May, 1917.

This article gives the derivation of the scale.

C. United States History

Barr's Diagnostic Tests in United States History. A test to diagnose historical abilities.

Address: A. S. Barr, 19 S. La Salle Street, Chicago, Illinois.

Harlan's Test of Information in American History. This is an information test in American history. It consists of ten exercises based on an analytical study of textbooks in American history. This test gives the teacher a measure of the efficiency with which the information side of American history has been taught. It may be used in first year of high school.

Publisher: Bureau of Educational Research, University of Illinois, Urbana, Illinois.

Sackett's Scale in United States History. This scale is composed of seven different tests based on carefully selected information in United States history. This scale can be used with pupils who have studied history above the sixth grade. It is difficult enough to test most high-school and college students. It is based upon the work of Bell and McCollum.

Publisher: L. W. Sackett, University of Texas, Austin, Texas. Also obtainable from Bureau of Educational Research, University of Illinois, Urbana, Illinois.

References: Bell, J. C. and McCollum, D. F. "A study of the attainments of pupils in United States history," *Journal of Educational Psychology*, 8:257-74, May, 1917.

Sackett, L. W. "A scale in United States history," *Journal of Educational Psychology*, 10:345-49, September, 1919.

Starch's American History Test, Series A. This test consists of 69 mutilated sentences in which the pupil is asked to write the missing words or phrases.

Publisher: University Cooperative Co., Madison, Wisconsin.

VIII. MATHEMATICS

A. General

Rogers' Mathematics Tests. These are a series of tests designed to measure several types of ability in the field of mathematics. The series include tests on arithmetic and algebra as well as geometry.

Publisher: Bureau of Publications, Teachers College, Columbia University, New York City.

Reference: Rogers, A. L. *Experimental tests of mathematical ability and their prognostic value.* (Teachers College Contributions to Education, No. 89.) New York: Teachers College, Columbia University, 1918.

Miss Rogers discusses in her thesis (1) summary of previous work; (2) general conditions of the present investigation, application of the tests and system of scoring; (3) analysis of mathematical ability; (4) prognosis of mathematical ability.

B. Arithmetic

Buckingham's Scale for Problems in Arithmetic. This scale consists of three divisions called: Division 1, Division 2, and Division 3, which are designed for different grades. Each division consists of ten problems evaluated in terms of difficulty. Difficulty is understood to mean the difficulty of getting the correct answer. Problems in Division 3 are in most cases sufficiently difficult for use in the high school.

Publisher: Bureau of Educational Research, University of Illinois, Urbana, Illinois.

Cleveland Survey Arithmetic Tests. They were designed for use in the survey of the Cleveland public schools. The series consists of fifteen tests, including four in addition, two in subtraction, three in multiplication, four in division, and two in addition and subtraction of common fractions.

Publisher: Bureau of Educational Research, University of Illinois, Urbana, Illinois.

References: Counts, G. S. *Arithmetic tests and studies in the psychology of arithmetic*. (Supplementary Educational Monographs, v. 1, no. 2) Chicago: University of Chicago Press, 1917.

Courtis, S. A. *The Gary public schools: measurement of classroom products*. New York: General Education Board, p. 147.

Courtis Standard Research Tests, Series B. These are the well-known Courtis arithmetic tests. The series consists of four tests, one on each of the operations with integers.

Publisher: S. A. Courtis, 82 Eliot St., Detroit, Michigan. Also obtainable from Bureau of Educational Research, University of Illinois, Urbana, Illinois.

References: *School report of the town of Southington, Connecticut*, 1919, p. 24.

An account of the use of the Courtis Arithmetic Tests, Series B, in first-year high-school commercial classes.

Courtis, S. A. *The Gary public schools: measurement of classroom products*. New York: General Education Board, 1919, p. 147 ff.

In addition to reporting on the use of these tests in the high school, the author gives an account of how his tests were constructed and a discussion of their validity and reliability.

C. Algebra

Coleman's Scale for Testing Ability in Algebra. This test consists of a series of exercises arranged in order of difficulty.

Publisher: W. H. Coleman, Bertrand, Nebraska.

Hotz's First-Year Algebra Scales. These scales consist of five lists of algebra exercises: (1) addition and subtraction; (2) multiplication and division; (3) equation and formula; (4) problems; and (5) graphs. In each list the exercises are arranged in order of difficulty. These scales are designed to determine the ability of pupils in terms of the exercises which just 50 percent of the class are able to solve correctly. They are instruments for measuring the abilities of classes rather than of individuals. No measure of rate of work is secured.

Publisher: Bureau of Publications, Teachers College, Columbia University, New York City. Also obtainable from Bureau of Educational Research, University of Illinois, Urbana, Illinois.

References: Hotz, H. G. *First-year algebra scales*. (Teachers College Contributions to Education, No. 90.) New York: Teachers College, Columbia University, 1918.

Dr. Hotz in his thesis discusses the origin and use of the First-Year Algebra Scales.

Cawl, F. R. "Practical uses of an algebra standard scale," *School and Society*, 10:88-90, July 19, 1919.

An account of the ways in which the Hotz First-Year Algebra Scales are of value to the superintendent and teacher.

Indiana Algebra Tests. A series of twelve tests which incorporated Monroe's Standard Research Tests in Algebra and six tests devised by H. G. Childs of Indiana University. Not available for use.

Reference: Childs, H. G. "The measurement of achievement in algebra," *Third Conference on Educational Measurements*. (Bulletin of the Extension Division, Indiana University, v. 2, no. 6) 1917, p. 171-83.

Sample exercises from tests are reproduced and results tabulated. Monroe's and Rugg's tests are compared.

Los Angeles Test in 9A Algebra. This test was devised by a committee of Los Angeles high-school teachers. It is based on the work of Rugg and Clark.

Reference: *Second year book of the Division of Educational Research, Los Angeles, California*, July, 1919, p. 113.

The test is reproduced with answers, results and graphs.

Monroe's Standard Research Tests in Algebra. These tests are based upon the simple equation and the steps involved in its solution. There are five tests, one upon each of the processes involved and a sixth test requiring a complete solution of a simple equation.

Publisher: Bureau of Educational Measurements and Standards, Kansas State Normal School, Emporia, Kansas.

References: Monroe, Walter S. "Test of attainment of first-year high school students in algebra," *School Review*, 23:159-71, March, 1915.

This contains an account of the derivation of the tests and their use with high school pupils.

Monroe, Walter S. "Measurement of certain algebraical abilities," *School and Society*, 1:393-5, March 13, 1915.

Second and third annual reports of the Bureau of Educational Measurements and Standards, 1915-1917, p. 76 ff.

Results for twenty-five cities are given.

Rugg and Clark's Standardized Tests in First-Year Algebra. This series includes sixteen tests which are intended to measure the ability to do all of the types of exercises in the work of the first year.

Publisher: University of Chicago Press, Chicago, Illinois.

References: Rugg, H. G. "The experimental determination of standards in the first-year algebra," *School Review*, 24:37-66, January, 1916.

The author states, "A report of the mathematics section of the Illinois High School Conference of a preliminary investigation for the purpose of: (1) establishing certain methods by which standards for measuring the outcomes of a year's instruction in high-school algebra may be constructed; (2) stating tentative results obtained in eight Illinois high schools; (3) making certain criticisms of the learning and teaching process in algebra."

Rugg, H. O. and Clark, J. R. "The improvement in ability in the use of the formal operations of algebra by means of formal practice exercises," *School Review*, 25:546-54, October, 1917.

Suggestions and directions for the use of such material in a cooperative experiment in the teaching of first-year algebra, carried on with a number of progressive teachers seeking to improve the teaching of the subject.

Rugg, H. O. and Clark, J. R. "Standardized tests and the improvement of teaching in first-year algebra," *School Review*, 25:113-32, 196-213, February-March, 1917.

This final report made to the Illinois High School Conference includes an account of the derivation of their tests. A reproduction (including "verbal problems") and the standardization of the tests.

Rugg, H. O. and Clark, J. R. *Scientific method in the reconstruction of ninth-grade mathematics*. (Supplementary Educational Monographs, v. 2, no. 1, April, 1918), p. 52-86. Chicago: University of Chicago Press, 1918.

"In this monograph the authors discuss the design and the construction of standardized tests in secondary mathematics and what the standardized tests revealed."

(To be continued in the April Number)

News Items and Communications

**Department of
Educational
Measurements,
Little Rock,
Arkansas**

A few principals and some teachers in the Little Rock public schools have in previous years shown interest in applying scientific methods to school work. This year, however, marks the beginning of organized effort. Superintendent R. C. Hall has been keenly interested in this new phase of education for some time, but the great need of a measuring department was not fully realized until September 12, 1919, at which time the writer was named as Supervisor of the Department of Educational Measurements.

The first and most difficult problems arising in the organization of the new department was that of securing an adequate number of assistants to carry on the work. Although the efforts of the department for the first year were to be confined to the fourth, fifth, and sixth grades, the children in these grades occupy 72 rooms and are divided into 92 groups. It was readily seen that the supervisor, who is principal of one of the grammar schools, could do but little of the actual testing and scoring of papers.

Relief came through the organization of a normal class at the Model School. This class is composed of twenty high-school graduates, under the efficient supervision of Miss Lulu B. Chase.

The giving and scoring of tests is conducted in the following manner: A conference is first held between the supervisor of the subject in which the test is to be given, the director of tests (Miss Chase), and the Supervisor of the Department of Educational Measurements. At this conference the test is selected. Miss Chase gives the normal class a course of training in methods of giving the test. When the students are sufficiently trained, they give the test in the various schools. The principals and teachers give the right of way to these examiners, making it possible usually to complete the testing in one day. The test papers are carried back to the Model School, where they are scored by the students and the results tabulated under the supervision of Miss Chase.

During the present school year tests have been given in reading and language. The results have been tabulated and returned to the schools. The principals and teachers of these subjects are doing very good work in studying these results, noting failures, seeking causes, and devising means of eliminating defects. We believe that a very great benefit is also being derived from this work by the members of the normal class. These students will enter the schools next year as teachers. They will be familiar with the methods of giving, scoring, and interpreting the results of standardized tests—a familiarity which will be expected of all teachers in the near future.

Now that work in educational measurement is passing rapidly from the universities to the public schools throughout the country, many superintendents are looking about for some plan by which departments of measurements can be put into operation in their school systems. We believe it a timely suggestion that the educational journals ask for and publish reports of this type from many places where successful plans are in operation. The above plan may not exactly fit any other system, but it is submitted in the hope that helpful suggestions may be gleaned from it.

G. T. HUCKABY

Supervisor of Educational Measurements

The JOURNAL OF EDUCATIONAL RESEARCH has received from Superintendent H. O. Dietrich of Kane, Pennsylvania, a report entitled "Findings, Statistics, and Recommendations of Reading Status of the Kane Schools as Revealed by the Administration of Monroe's Standardized Silent Reading Tests, December, 1919."

**Use of Reading
Tests at
Kane, Pa.**

The editor of this department wishes to take this opportunity to say that he is glad to receive such reports and to invite other school men to be kind enough to send a copy of such reports as they may issue in either printed or manuscript form.

The limitations of space prevent the publication in full of the report submitted by Superintendent Dietrich. Since the value of an educational test depends upon the constructive changes which are made upon the basis of the information revealed by the tests, his recommendations are of especial interest and for that reason they are reproduced below.

RECOMMENDATIONS

"1. That variability be reduced. It is too large. The only variability plausible is that of grade seven in rate and grades five, seven, and eight in comprehension. Variability between grades and buildings is also too great.

"2. That the rate of silent reading be increased up to and including grade six.

"3. That comprehension be stressed in all grades.

"4. That an effort be made to reduce overlapping of grades.

"5. That mechanics be stressed in the first three grades so that the rate of silent reading will be increased.

"6. That the quick perception and sense content method be used in preference to the phonetic in the primary reading.

"7. That all reading in intermediate grades shall be done under guidance of specific purposes.

"8. That practice in enlarging the span of attention should be given in the lower grades.

"9. That comprehension can be increased by insuring a sound mechanical foundation.

"10. That the teachers prepare reading exercises for the pupils which are not found in regular readers.

"11. That the teachers conscientiously study the charts, graphs, and statistics, and make reports so that they may better be able to study each child's case.

"12. That the individual scores of each child be kept on file so that the needy cases may receive attention immediately upon diagnosis of their trouble.

"13. That the individual defects or impediments be noted for each case studied.

"14. That the subject of reading be made a constructive study on the part of all teachers."

In common with other cities, Boston has faced the problem of increasing the salaries of teachers in public schools. The school committee has recently published the results of an investigation by Superintendent Frank B. Thompson concerning the "Salaries of public school teachers in cities of over two hundred thousand population in the United States and in selected cities and towns in Massachusetts."¹ From a supplementary publication we quote the following:

PRINCIPLES UNDERLYING PROPOSED SALARY INCREASES

"The proposed new salary schedule has been prepared by the School Committee on the basis of three underlying principles which are fundamental:

1. Largest proposed increase in salary goes to the lowest paid teachers in the service.

2. Some increase in salary is proposed for everyone in the employ of the School Committee.

3. The amount of proposed increases and the readjustments conform as closely as possible to the supply of teachers.

"The lowest paid teachers in the Boston service are the kindergartners and the elementary assistants. Accordingly, they have been given the largest increase, both in actual amount and in percent of increase. No other rank of teacher is given more.

"Some increase in salary has been provided for everyone. The nurses, school physicians, the attendance officers, and the directors and supervisors have all been allowed increases in salaries in the new schedule. In addition, the evening school teachers and the teachers in summer review schools will hereafter receive more money for their services. Substitutes, emergency assistants, employees in the school centers, and clerks in the executive offices have not been left out of consideration. A substantial increase is also proposed for school janitors. In short, the School Committee has consistently followed the principle of something for everybody in the new schedule.

"In the varying amounts of proposed increases for different ranks of teachers, the School Committee has recognized that there are inequalities

¹ School Document No. 19, 1919, Boston Public Schools, November, 1919.

in the present schedule, as pointed out by teachers, and has tried to readjust some of them. In addition, the School Committee has also taken into account the salary schedules, both present and proposed, in other cities. Assuming for the moment that living conditions in several cities may be identical, teachers are not likely to leave Boston to teach in other cities if Boston salaries are equal, or about equal, to salaries in those cities. Since comparatively few teachers who resign from teaching go into other occupations, the probability of losing teachers to other cities depends on the salary there. In establishing the proposed new salaries the School Committee has fixed them high enough so that the salaries of teachers elsewhere will not be attractive to Boston teachers. Whether the proposed salaries are high enough to attract into the profession those qualified by nature and training to teach is a question which the Committee has not undertaken at this time to determine fully and conclusively. It is, perhaps, significant that colleges and universities generally, which for the last few years have suffered from a falling off in attendance, are now, in some cases, almost overwhelmed by the number of students, some of whom will eventually take up the teaching profession, and will, undoubtedly, be attracted thereto by the present general movement to increase teachers' salaries."

National Association of Directors of Educational Research

(E. J. ASHBAUGH, *Secretary and Editor*)

Well, what do *you* think of it? Do you appreciate the fact that the new JOURNAL OF EDUCATIONAL RESEARCH, the journal you now hold in your hands, is our journal, your journal? Do you feel a personal pride in it and a personal responsibility for the material that goes into it; for the influence which it wields? Do you know that the first issue had not been out a week when it brought to your secretary an inquiry concerning membership in the National Association of Directors of Educational Research?

These are questions to which we wish every member of the Association would make answer for himself. Say it aloud. See and hear how it sounds and then remember that you have covenanted with yourself to give something of your time and thought to this publication. We are pledged to the idea of organized bureaus of research in city school systems, state departments of education, and colleges and universities. We believe that such organization is the most efficient means of gathering and evaluating the data which are necessary for progress in public education. Each bureau is working upon problems peculiar to itself, but many of the data gathered are of more than local interest. Let each member feel morally obligated to send to Dr. Buckingham, editor-in-chief, or to your secretary within the next thirty days, a report on some work he has done. The report need not be long. It should be clearly written, presenting a definite problem, concisely stating the method of procedure and the results. It should be closely allied with the work of public school teachers, administrators and supervisors. It should give an answer to some question of importance to them or show a way of attack which they can use. We are sure you have the material for such a report. Will you take the time to put it into shape for publication and send it in? By so doing you will at once discharge at least a part of your responsibility, will guarantee the success of the Journal, and be of real service to education throughout the country.

What are people thinking of the new JOURNAL OF EDUCATIONAL RESEARCH? We believe it will be quite in order to quote from some of the letters which reached our desk within a few days after the first number was issued.

"I wish to say that no educational paper has ever appealed to me as much as the first copy of the Journal of Educational Research."—An Ohio superintendent of schools.

"The Journal of Educational Research looks like the best bet in the country."—Michigan.

"I have just received the Journal of Educational Research. It is going to give us just what we need."—Iowa.

These and other communications indicate that those who have seen it, like it. But has it occurred to you that one or more of your progressive, wide-awake friends may not have seen it yet? Could you do them a greater service than to call it to their attention? I want every friend of mine in Iowa to get acquainted with it. What about you in your state?

Director J. C. De Voss of the Kansas State Normal School has sent in the following news notes. If you are interested in any of the problems upon which he is working, write him and thus coordinate your work and get greater results.

"We have given the Army Intelligence Test Alpha to the Freshmen and Sophomores of the Kansas State Normal School and to high school students of several junior high schools, also to some pupils in the intermediate grades. We are preparing a manual to be used with the Army Alpha tests in the public schools. We are also preparing age and grade distributions of the scores obtained. We are standardizing a new Latin Derivative test and are making good progress in preparing a scale for use in Wood-working in Manual Training.

"One very interesting feature of our work this year is that of extension classes among teachers in five cities. I believe this work is sure to grow in volume and in importance in Kansas."

Dr. Ashbaugh and Dr. Greene of the Bureau of Educational Service, University of Iowa, have been engaged in two very interesting surveys this winter. The one was concerned with the course of study in the junior high school grades and involved the comparison of the relative efficiency of teaching Latin and Algebra in the eighth and ninth grades. The other was the formulation of a constructive building program for a school district in which the financial possibilities were only about one-third of the immediate needs. This latter study involved a comparison of the assessed valuation of real estate and the actual value as shown by sale price. If any reader wishes some real enlightenment at the expenditure of a couple of hours' work, it is suggested that he go to the county recorder's office and list the sale price of a number of pieces of property and then compare these with the values shown on the assessor's books in the auditor's office.

Journal of Educational Research

Editorial Office, Urbana, Illinois

Editor-in-chief B. R. Buckingham, University of Illinois

VOLUME I

APRIL, 1920

NUMBER 4

CONSTRUCTING A LANGUAGE AND GRAMMAR SCALE

W. W. CHARTERS

Carnegie Institute of Technology

INTRODUCTION

In our studies of the efficiency of school children in using correct speech, two difficulties were recognized at almost the beginning of the investigation. In the correction of language forms upon a language level without knowledge of grammatical rules and definitions, it was found that few teachers had definite information about the errors which their children made. They recognized that errors occurred and that in a general way some were more common than others; but any detailed knowledge of exactly how such errors should be attacked was almost entirely lacking. This lack of definite information prohibited them from making local attacks upon specific errors; and this is, in almost all cases, with some marked exceptions, true not only of the individual teacher but also of whole school systems. For, the tendency is to take pot shots at errors and this in spite of the fact that the only method of correction is to attack them one at a time and that a certain number of errors should be assigned to each grade for frontal attack with the supporting assistance of the grades above and below.

It was found also, that the study of grammar seemed to have an almost negligible effect upon the correction of errors. A steady increase of skill from grade to grade is noticeable before grammar is studied but in the grades where grammar is taught and where a rapid increase in skill is expected there did not seem to be a greater increase than in the lower grades in which instruc-

tion is not given. If this is true, then two results must follow. Either grammar should be eliminated as a grade subject of study or methods should be used by which it will have a noticeable effect upon language usage. Very serious doubts as to its usefulness in grades are growing rapidly. But until it is clearly demonstrated that it does not and cannot affect the correction of errors, the doubt will not be resolved.

The probable reason for its failure to function as a basis of correction is that the educational world has taken the indefensible position that incorrect forms must not be placed before children. That the verb must agree with its subject in person and number is taught in connection with the analysis of correct sentences only. For in select educational circles it is considered bad form to place on the board sentences such as, "Was you looking for something?" or "He don't look well" and ask children to apply the rule to their correction. This is probably caused by a confusion of thought. It is entirely defensible to argue that the incorrect form should not be made conspicuous when memory alone is involved. Before the study of the grammatical rule which is found broken the pupil may become confused by seeing both the correct and incorrect form given equal prominence; but once the rule has been taught and memory is reinforced by reason, there is every reason for fusing the error and the rule into one experience. Only by this process, carried on in the school room, will the pupil recognize that rules have any use as correctives. Until that is done deliberately and thoroughly, grammar will be concerned merely with the analysis of correct sentences, and errors will live in a world apart. And if this is not done, grammar must disappear from the grades. It is difficult at its easiest, and because it is difficult, it leads to confusion and probably, as mental exercise, does more harm than good. If it can demonstrate its utility as a corrective, it may hope to live. Like other vested interests it must demonstrate its ability to render service to the public.

Following the realization of the two facts that language errors are not attacked methodically and persistently and that grammar does not seem to function, the idea was developed that formal tests might be devised which would place information upon both these points at the disposal of the teacher. Such tests should obviously be diagnostic and should display the ability of the

pupils to recognize incorrect forms, to correct them and to give the reason for correction where this depended upon rules of grammar already learned.

METHOD

By language tests are meant those in which the pupil is expected merely to recognize the correct form and by grammar tests are meant those in which he can provide in addition, the basic rule for correction couched in grammatical terminology. To say that "He gave it to me and Mary" is incorrect because it is "impolite" or does not "sound right" is not a use of grammar. The grammatical rule states that the first personal pronoun stands last in a series; for only in this latter case is the terminology of grammar used.

In constructing the language tests several alternatives presented themselves. The first form considered consisted of the use of alternative forms, the incorrect one of which was to be crossed out, as "*(Was) (Were) you looking for something?*" This form could be quickly treated by the pupils and rapidly scored by the teacher. But it would not give an accurate display of the normal ability of the pupils to use the correct form. For normally, when an individual is about to ask the question, "*(Was) (Were) you looking for something?*" he does not have the two forms in his field of attention; if they were both there the hint would be sufficient to supplement his memory in the choice of the correct one.

This form was discarded and the device accepted of stating, "In the following sentences some are correct and some are incorrect. On the line beneath the sentence write the correct form." It was felt that the pupil would thereby have to make a complete and unassisted decision in a situation approximating very closely to the conditions of normal speech.

In the grammar test the same sentences are used but in addition space is provided on the right half of the test page for giving the reasons for any corrections made. The form is shown in the following portion of one of the tests. The complete test has forty items. The title and the provisions for general information about the class and the pupil are omitted.

GENERAL INSTRUCTIONS

After telling the children not to turn the papers over, distribute the folders, this side up. Have each pupil fill in the blanks at the top. Then say: "Look at the directions while I read them to you."

DIRECTIONS: (To be read by the teacher and pupils together)

"This test is given to pupils who have studied language and grammar to see how well they are able to tell when sentences are right and when they are wrong. Now look at the sample below."

1. I told him to go. {
 }

"The plan is to read this sentence over carefully and see if it is right. If it is right make a cross on the dotted line below the sentence. The sentence, 'I told him to go,' is right so we shall make a cross on the dotted line below it. Make the cross now." (The teacher should pause to allow time for doing this and should pass around the room in the lower grades to make certain every pupil understands.)

"If the sentence is not right we are to put the words on the dotted line below it. Let us try one that is not right."

2. I told she to go. {
 }

"The right form is 'I told her to go.' So we shall write 'her' on the line below it."

"On the right-hand side we are to tell why we make the change on the left-hand side. We always begin our reason for the change with the word *Because*. In this case we write—*Because 'her' is in the objective case, object of the verb, told.* (The teacher should pause until the pupils write the reason at the right of the sentence.) We must not write too large and we may use all of the lines on the right-hand side so that we may have plenty of room for giving our reasons."

"We shall now try a third exercise."

3. It is hisn. {
 }

"The right form is *It is his*. We make the change *Because his is the possessive form of the pronoun, he*. Fill this in correctly, writing 'his' on the line below and beginning the reason on the right-hand side with the word *Because*. Do this now."

"Remember that if the sentence on the left is all right, we do not need to write anything on the right-hand side because we have to make no change and so have to give no reason for making a change."

The teacher should pass around the room and see that the children fill these in correctly and when all have finished he should say, "Begin work on the exercises below. Stop when you have finished this page and bring your paper to me." This is not a speed test so ample time should be allowed to each to finish. This test should be completed in two periods during the same day. Part I should occupy the first period, and Part II the second. At the

beginning of the second period, go over the above instructions so far as may be necessary before directing the children to begin on Part II. After the children have finished Part II, collect the papers.

| | |
|--|----------|
| 1. John went to town. | 1 |
| | |
| 2. May Inez and me go? | 2 |
| | |
| 3. It teaches a person something you may use. | 3 |
| | |
| 4. Yous girls play too hard. | 4 |
| | |
| 5. When one lives in town they hear noises. | 5 |
| | |
| 6. That is hern. | 6 |
| | |
| 7. Who do you want? | 7 |
| | |
| 8. Him and I will do it. | 8 |
| | |
| 9. Us boys did it. | 9 |
| | |
| 10. The boy was raised by its own mother. | 10 |
| | |
| 11. These books are yourn. | 11 |
| | |
| 12. It was a book with riddles in them. | 12 |
| | |
| 13. Are those them? | 13 |
| | |

CONTENT OF THE TESTS

From studies made in Kansas City, Detroit, and elsewhere, a rather complete list of common grammatical errors in the verba-

tim expressions of the children was obtained. These were used and selection was made upon the basis of frequency of occurrence. The following tests were constructed in duplicate forms.

I. LANGUAGE

Pronouns

Verbs A

Verbs B

Miscellaneous

II. LANGUAGE AND GRAMMAR

Pronouns

Verbs A

.....

Miscellaneous

Each test has forty items, one or two of which are correct, to keep faith with the children in conforming to the promise that they would find that some of them were correct. The number, forty, was determined partly by the size of practicable pages and partly by the fact that it gives a comprehensive list of the errors at hand. The forty do not include all items as do the tests of the fundamental operations in arithmetic, but all frequently occurring and many infrequently noticed errors are included so that for purposes of class instruction, where only the commoner ones will be studied, as distinguished from individual errors made by only one or two, the significant ones are undoubtedly included. Some pupils will make other errors but all of sufficient frequency to receive class attention are present in the tests.

For purposes of comparison of efficiency at different times two forms of the test (Form I and Form II) are available. The forms are alike in that the same types of errors are included in each and in the same order but differ in the sentences in which the errors are found.

The language tests are constructed so as to be given from grades III to VIII or in the high school. The grammar tests will naturally be given only in those grades which studied the grammatical facts involved.

TIME LIMIT

In giving the tests no time limit is prescribed at present. This has been omitted because the sentences necessarily vary in length if approximately natural sentences are used; and relative ability, as shown by the test, would be greatly affected by speed in handwriting. In the grammar division of the test time conditions are further disturbed by the fact that no standardized form for stating grammatical rules has been accepted or proposed by the

schools. In addition to these considerations opposing the setting of a time limit is the further fact that the test being diagnostic in character it is essential that all the errors should be attacked by each pupil.

However, in our most recent modification of the language section of the tests we are experimenting with the writing of corrections by asking the pupils not to write the *whole sentence* in correct form but to insert, on the line beneath, the correct *words*.¹ This may make it possible to set a time limit on the language section; but it does not remove any difficulty in the grammar section.

WEIGHTING

A very careful study was made of approximately three thousand one hundred seventh and eighth-grade test papers on pronouns to determine the weight of each item. The plan followed was to determine the number and percent of incorrect replies and using sigma (σ) as a basis and 2.5σ as zero to give the approximate weight to each. Fortunately, it was found that the relative ranking of individuals was not appreciably affected by dropping the two decimals from the weighted amount and using the nearest integer. Furthermore, it was discovered that the correlation between rankings determined upon the bases of unequal and equal weights for each item was slightly over 0.90. Consequently, it was decided to adopt the simple plan of counting each item as possessing unitary value.

It is entirely possible that if zero ability were more closely defined than 2.5σ and if it were based on zero third-grade ability rather than upon zero seventh-grade skill, as was done in our study of the three thousand one hundred cases, the weighting might be more significant than has been found in the latter case. But for diagnostic purposes, which is the professed function of these tests, this consideration is not of great significance.

What was true of the weighting of language skill when based on percentages of failure was more noticeable in grammar weightings because of the very large percentages of failure to give reason for correction. The three thousand one hundred papers examined had a median percentage incorrect of over fifty.

¹ The directions reproduced on page 252 above are the ones used in this modification of the original test.

SCORING

The scoring of the language section presents no serious difficulty since the forms tested are those upon which any well-informed teacher is competent to judge. However, scoring cards are provided to expedite the work.

In the grammar section the absence of standardized forms, as in addition or multiplication, presents a rather cumbersome difficulty. For instance, the sentence "Him and I will do it" was written correctly as "He and I will do it." The reasons given for changing to the correct form, *He*, were the following among many.

- (1) Because *He* is in the nominative case, subject of the verb *will do*.
- (2) Because *He* is the subject of the sentence.
- (3) Because *He* is in the nominative form.
- (4) Because *He* is subjunctive form.
- (5) Because it would not be right to say *Him will do it*.
- (6) Because we should use *He* instead of *him*.
- (7) Because *He* is the noun.
- (8) Because it seems to be turned around.
- (9) Because *me* is objective form, object of the verb *and*.

It is obvious that (7), (8), and (9) are absurd and that (5) and (6) give no grammatical basis. It is also clear that (1) is a complete answer consisting of both the description of the case (nominative) and the syntactical relation governing it (subject of the verb "will do"). The difficulties of the decision of all such cases is seen in (2), (3), and (4). This difficulty arises from the fact that they each include one of the two parts of the complete answer. The second (2) gives the syntactical relation; (3) gives the case form and (4) is, probably, a mixture of (2) and (3).

It was decided that while teachers should naturally expect a pupil to give both elements as a standard of class instructions, an examiner should rate the answer as correct if one or the other of the two elements were given. This would give a more accurate gauge of ability than would the elimination of (2), (3), and (4). For there is no standard form for the answer and if the pupil knows the one element he probably knows the second since they are interdependent, and in any case either one of the two elements affords a good clue for practical speech.

The norms to be established for these scales will include not only the median but, in addition, the 25 and 75 percentile. This is necessary in any diagnostic scale. For since it focuses attention on the individual, it is obvious that no stimulation will be derived from its use by the half of the class above the median if only the median is used as a standard.

THE USE OF THE TESTS

These tests are designed to serve four purposes. In the first place, a teacher who gives one of them to his class and has them marked and tabulated has a clear idea of the points at which to work in the acquiring of language skill, an idea much clearer than that afforded by incidental and more or less casual observation. In the second place, it is possible by the use of the alternative forms to see what progress is made during any period. In the third place, they provide a means of discovering whether or not grammar is of any use in the control of errors, as can be easily found by testing freshman students in high school who have and have not studied grammar in the grades. Finally it presents ocular proof to individual pupils that they do not know how to use certain forms with accuracy and ease.

These tests are still in tentative form and the cooperation of those interested is invited for assistance in smoothing out the defects and in establishing standards for the grades and the high school. Copies of the tests and other information may be obtained from the Bureau of Educational Research at the University of Illinois, Urbana, Illinois.

PSYCHOLOGICAL TESTS AS A MEANS OF MEASURING THE PROBABLE SCHOOL SUCCESS OF HIGH-SCHOOL PUPILS

W. M. PROCTOR

Leland Stanford Junior University

The validity of the Stanford-Binet Scale, when applied to high-school pupils, has already been discussed by the writer in the issues of *School and Society* appearing October 19, and 26, 1918.¹ In those articles it was shown that very significant correlations had been obtained between intelligence quotients,² (I. Q.'s), resulting from the individual tests of 107 high-school pupils and the school marks earned by the same pupils during the school year, 1916-1917; also between I. Q.'s and teachers' estimates of intelligence made during the same year.

Two years and a half later there were 66 of the original 107 high-school pupils remaining. Teachers who had known all of these pupils during their stay in the high school were asked to give estimates of their intelligence upon the same rating sheet as that which was used in 1916-1917. All school marks earned during the two and one-half years were averaged. Correlations were then found (a) between the I. Q.'s obtained in 1916-1917 and the teachers' estimates made in 1919; (b) between the average of all school marks earned up to April 1, 1919 and I. Q.'s obtained in 1916-1917; and (c) between the average school marks and the teachers' estimates made in 1919. Table I shows the close agreement between the correlations obtained in 1916-1917 and those found in 1919.

Table I shows that the correlations obtained in 1918-1919, when the same comparisons were made as in 1916-1917, were

¹ Proctor, W. M. "The use of intelligence tests in the educational guidance of high-school pupils," *School and Society*, 8:473-78, 502-9, October, 19-26, 1918.

² The intelligence quotient is obtained by dividing the mental age by the chronological age. Thus a twelve-year-old, chronologically who tested eight years old mentally would have an I. Q. of 0.66, expressed for convenience "66." The I. Q. is an index of relative brightness. For further discussion of intelligence quotient see: Terman, L. M. *The intelligence of school children*, Houghton Mifflin and Co., 1919, p. 7.

TABLE I. COMPARISON OF CORRELATION OBTAINED
IN 1916-1917 AND THOSE OBTAINED IN 1918-1919^a

| Year | Correlation between I. Q. and Teacher Estimates | Correlation between I. Q. and School Marks | Correlation between School Marks and Teacher Estimates | Total Number of Cases |
|--------------|--|---|--|--------------------------------|
| 1 | 2 | 3 | 4 | 5 |
| 1916-17..... | 0.586 ± 0.043 | 0.545 ± 0.046 | 0.702 ± 0.033 | 107 |
| 1918-19..... | 0.583 ± 0.055 | 0.487 ± 0.063 | 0.667 ± 0.046 | 66 |

^a Pearson's formula³ (shorter method) was used in making all correlations.

practically as high as those obtained in the first instance. The results of two and one-half years of follow-up work seem to indicate that the person who made the original tests with the Stanford-Binet Scale in 1916-1917, would have been in a position to give very helpful advice to all of the pupils tested by him; also that his predictions as to the possible educational future of each of these pupils would have deserved serious consideration by parents and teachers.

As a means of discovering individual differences between school children in order that they may be grouped in classes according to ability, the individual psychological test has been shown to be a helpful tool. From the standpoint of school administration, however, the individual test presents serious difficulties. The time required to give an individual test to a high-school pupil varies from 40 minutes to 120 minutes. The total number of pupils that can be examined by a single examiner in a day will seldom exceed ten. The use of the Stanford-Binet abbreviated scale enables an examiner to test from 15 to 25 pupils in a day. Even so, it is impossible to use the individual method when a rapid survey of an entire school population is to be undertaken.

Group mental examinations afford the only means of meeting the demand for a speedy and reliable method of measuring the mental abilities of large groups of people. Under the supervision

³ Rugg, H. O. *Statistical methods applied to education*. New York: Houghton Mifflin and Co., 1917, p. 274.

of Dr. L. M. Terman the writer directed the giving of Examination *a* (Form A) and Examination Alpha (Form 5) of the Army Scale to 1,349 high-school pupils, representing eight California high schools, during the school year 1917-1918.

Examination *a* consists of ten tests: (1) Oral Directions, (2) Memory for Digits, (3) Disarranged Sentences, (4) Arithmetical Reasoning, (5) Information, (6) Synonym-Antonym, (7) Common Sense, (8) Number Series Completion, (9) Analogies, (10) Number Comparison. The total possible raw score is 237. This test was given to portions of the first-year high-school classes in the Oakland Technical, Oakland Central and Fremont, Oakland, high schools and to all of the pupils present on the day of the examination at the Palo Alto Union High School. The total number of pupils was 715.

Group Examination Alpha consists of eight tests: (1) Oral Directions, (2) Arithmetical Reasoning, (3) Practical Judgment, (4) Synonym-Antonym, (5) Disarranged Sentences, (6) Number Series Completion, (7) Analogies, (8) Information. This test was given to all pupils present on the day of examination in the San Mateo, Redwood City, Mountain View, and Santa Clara union high schools. The total number of these pupils was 624.

These group mental examinations were applied to all of the pupils in each of the high schools enumerated above at exactly the same time. A sufficient number of examiners, trained by Dr. Terman, were taken to each high school, to cover the entire high school in one forty-five minute period. The size of the groups ranged from 40 to 150. It took the writer and his assistant a total of 134 hours to test 107 high-school pupils by the individual method. Six trained examiners were able to give Examination *a* to 350 Palo Alto high-school pupils in 45 minutes. The test blanks were scored by university students. Their work was carefully checked and the results tabulated by the writer.

I. GROUP TEST RESULTS COMPARED WITH INDIVIDUAL TEST RESULTS

One hundred and sixteen of the high-school pupils tested with Examination *a* had previously been given the Stanford-Binet Scale. Table II makes comparison of the two kinds of mental examination. Although no Binet I. Q.'s are found in the group 140-149, six Army Scale I. Q.'s are between 140 and 149. This

is because a higher mental age is attainable on the Army Scale than on the Binet Scale. For example, a high-school boy fifteen years and two months old passed every test in the Stanford-Binet

TABLE II. RELATION BETWEEN THE I. Q.'S OF 116 HIGH-SCHOOL PUPILS EARNED ON THE STANFORD-BINET SCALE AND THE I. Q.'S OF THE SAME PUPILS EARNED ON EXAMINATION *a*, ARMY SCALE

| I. Q.'s on Stanford-Binet Scale | I. Q.'s on Examination <i>a</i> , Army Scale | | | | | | | Totals |
|---------------------------------|--|-------|---------|---------------------|---------|---------|---------|--------|
| | 80-89 | 90-99 | 100-109 | 110-119 (Median) | 120-129 | 130-139 | 140-149 | |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 130-139..... | | | | | 3 | 3 | 5 | 11 |
| 120-129..... | | | 1 | 3 | 7 | 9 | 1 | 21 |
| 110-119 (Median)..... | | 2 | 8 | 11 | 9 | 2 | | 32 |
| 100-109..... | | | 13 | 3 | 4 | | | 20 |
| 90-99..... | | 7 | 11 | 8 | | | | 26 |
| 80-89..... | | 5 | | | | | | 5 |
| 70-79..... | 1 | | | | | | | 1 |
| Totals..... | 1 | 14 | 33 | 25 | 23 | 14 | 6 | 116 |

Median for Binet I. Q.: group, 110-119

Median for Army Scale I. Q.: group, 110-119

Correlation, Pearson's formula, 0.736; P. E., 0.029

Scale, thus earning a mental age of nineteen years and six months, and an I. Q. of 129. On the Army Scale, Examination *a*, he made a raw score of 219, corresponding to a mental age score of twenty-one years and eleven months, and an I. Q. of 144.⁴ In other words, the Stanford-Binet Scale does not give the superior high-school pupil an opportunity to earn as high an I. Q. as he can earn on the Army Scale. This factor would tend to lower the

⁴ Mental age norms for both tests of the Army Scale were worked out by Dr. Samuel Kohn and the writer. It was found that about fifteen points on Examination *a* and twelve points of raw score on Examination Alpha corresponded roughly to a mental age year. Possible raw score, Examination *a*, 237, possible mental age twenty-three years; possible raw score, Alpha 212, possible mental age twenty-four years and nine months.

correlation between the two sets of I. Q.'s. The correlation obtained in Table II (+0.736) is a strong indication that if the Stanford-Binet Scale is a valid means of finding the mental level of high-school pupils, the Army Scale Examination *a* is also valid for the same purpose.

II. ARMY SCALE RESULTS COMPARED WITH HIGH-SCHOOL MARKS

The school work of all the high-school pupils examined by means of the two army group tests was carefully followed up for the school years 1917-1918 and 1918-1919. The marks given were reduced to a comparable basis by assigning arbitrary values to each type of mark employed by the various high schools. Thus an "A" or a "1" was given a value of 95; a "B" or a "2," a value of 85, etc. Letters or numbers with plus and minus signs were given intermediate values. All of the marks earned by a given pupil were averaged, but no case was included in the tables unless the marks for at least two semesters of school work, the equivalent of one year, were available.

1. *Army Scale, Group Examination a*.—Table III makes a comparison between the I. Q.'s obtained from the Group Examination *a* and the quality of high-school work of 494 high-school pupils. The total number taking the test was 715, but only 494 cases had ratings for one year of school work.

Since the correlation obtained (+0.343) is 12.8 times the indicated P. E., it has considerable significance. While it is not as high as the correlation between Binet I. Q.'s and school marks found in Table I, there are several factors which may have tended to lower the correlation. The army tests were designed for use with soldiers in cantonments. Many of the questions have to do with matters of common knowledge about a military camp, but with which high-school pupils have no acquaintance. This would especially affect the scores of high-school girls. There are twenty cases falling in I. Q. groups below 95, where the indicated school work is of a quality of 80 percent or above, and fifteen of these cases, or 75 percent, are girls.

Another factor which may have tended to lower the correlation is the skewness of the school marks curve toward the high percents. The median school marks group in Table III is 80-84.

TABLE III. CORRELATION BETWEEN THE I. Q.'S OF ARMY GROUP EXAMINATION *a* AND THE QUALITY OF SCHOOL WORK OF 494 HIGH-SCHOOL PUPILS

| School Marks | Army Group Examination <i>a</i> I. Q.'s | | | | | | | | | | Totals |
|--------------------|---|-------|-------|-------|---------|---------------------|---------|---------|---------|--------------|--------|
| | 84 or Lower | 85-89 | 90-94 | 95-99 | 100-104 | 105-109 (Median) | 110-114 | 115-119 | 120-124 | 125 or Above | |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| 90 or above | | | 3 | 2 | 2 | 3 | 6 | 6 | 6 | 6 | 34 |
| 85-89 | | | 5 | 9 | 18 | 24 | 22 | 24 | 8 | 4 | 114 |
| 80-84 (Median).... | 1 | 2 | 9 | 28 | 41 | 46 | 30 | 13 | 9 | 3 | 182 |
| 75-79 | 1 | 3 | 9 | 19 | 19 | 19 | 19 | 7 | 2 | | 98 |
| 70-74 | 2 | 3 | 4 | 14 | 11 | 11 | 7 | 1 | 1 | | 54 |
| 65-69 | | 1 | 1 | 4 | 1 | 1 | 2 | 1 | | | 11 |
| 55-64 | | | | | | 1 | | | | | |
| Totals | 4 | 9 | 31 | 76 | 92 | 105 | 86 | 52 | 26 | 13 | 494 |

Medians: I. Q.'s group, 105-109; school marks, group 80-84

Semi-interquartile ranges: I. Q.'s, 6 points; school marks, $4\frac{1}{2}$ percent

Correlation, Pearson's formula, 0.343; P. E., 0.027

This represents a grade of "B" or "2." It is necessary for these high-school pupils to receive marks of "B" or "2" in all subjects required for university recommendation. The Palo Alto high school is in close proximity to Stanford University, and the Oakland high schools are in the immediate vicinity of the University of California. The fact that 68.8 percent of the grades given to these 494 high-school pupils were "B" or above is an indication that teachers were influenced in their marking by the demand for "B" grades for university recommendation. Many pupils with just *average* ability were given marks ranking *superior*. Hence there would be a rather wide difference between their mental ability as shown by the tests and their school progress as shown by their marks.

2. *Army Scale, Group Examination Alpha*.—Table IV compares the I. Q.'s of 480 of the high-school pupils of San Mateo,

Redwood City, Mountain View, and Santa Clara with their school marks. There were 624 pupils belonging to these high schools who took Examination Alpha, but only 480 cases came under the rule requiring ratings for one whole year of school work.

TABLE IV. CORRELATION BETWEEN THE I. Q.'S OF THE ARMY GROUP EXAMINATION ALPHA AND THE QUALITY OF SCHOOL WORK OF 480 HIGH-SCHOOL PUPILS

| School Marks | Army Group Examination Alpha I. Q's. | | | | | | | | | | Totals |
|--------------------|--------------------------------------|-------|-------|-------|---------|---------------------|---------|---------|---------|--------------|--------|
| | 84 or Lower | 85-89 | 90-94 | 95-99 | 100-104 | 105-109 (Median) | 110-114 | 115-119 | 120-124 | 125 or Above | |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| 90 or over. | | | | 3 | 3 | 15 | 12 | 9 | 9 | 5 | 56 |
| 85-89. | | | | 8 | 17 | 15 | 24 | 13 | 6 | 6 | 89 |
| 80-84. | | | 4 | 6 | 22 | 21 | 20 | 10 | 5 | 1 | 89 |
| 75-79 (Median).... | | | 7 | 25 | 33 | 23 | 10 | 7 | 4 | | 109 |
| 70-74. | | 4 | 10 | 18 | 14 | 22 | 12 | 1 | 1 | | 82 |
| 65-69. | 1 | 3 | 3 | 12 | 7 | 8 | 8 | 1 | | | 43 |
| 55-64. | | | 2 | 5 | 3 | 1 | 1 | | | | 12 |
| Totals. | 1 | 7 | 26 | 77 | 99 | 105 | 87 | 41 | 25 | 12 | 480 |

Medians: I. Q.'s group 105-109; school marks, groups 75-79

Semi-interquartile ranges: I. Q.'s, 6 points; school marks, $6\frac{1}{2}$ percent
Correlation, Pearson's formula, 0.413; P. E., 0.026

The correlation obtained in Table IV (+0.413) is 15.9 times the indicated P. E. and 0.07 higher than the correlation found in Table III. The higher correlation found in this table may be due to the fact that Examination Alpha was the result of careful revision of the first series of tests in the light of preliminary experimentation in three army cantonments and the returns from the tests of several thousand school children. Also, the high schools in which Examination Alpha was applied were not in such close proximity to universities as were the high schools represented in Table III. The group median for school marks is 75-79, and only 48.8 percent of the ratings given in these four high schools ranged as high as "B" or over.

When account is taken of the possible reasons for the difference in the correlations as shown in Tables III and IV, it is safe to say that they are of approximately equal value. For use in the public schools Examination Alpha is the better scale, because it consists of but eight tests, can be given in less time, is scored more rapidly, and costs less to print.

III. INDIVIDUAL AND GROUP MENTAL TESTS AS MEANS OF INDICATING PROBABLE RETENTION OR ELIMINATION OF HIGH-SCHOOL PUPILS

1. *Elimination among pupils tested with the Stanford-Binet Scale.*—On the first day of April, 1919, it was found that 41 of the original 107 pupils tested in 1916-1917 with the Stanford-Binet Scale had dropped out of school, leaving 66 still in attendance. Table V gives the distribution of the 107 cases, showing the I. Q.'s of pupils who left high school to go to work, of pupils who were transferred to other high schools, and of pupils who still remain in the same high school.

TABLE V. DISTRIBUTION ON APRIL 1, 1919, OF 107 HIGH-SCHOOL PUPILS TESTED WITH STANFORD-BINET SCALE IN 1916-1917

| Binet I. Q.'s Earned in 1916-1917 | Total Number of Cases | Distribution April 1, 1919, by Percents | | |
|---|-----------------------------|---|---|---|
| | | Out at Work (%) | Out, Transfer'd to Other High Schools (%) | Remaining in Same High School (%) |
| 1 | 2 | 3 | 4 | 5 |
| 79 or lower..... | 1 | 100 | 0 | 0 |
| 80-89..... | 7 | 72 | 14 | 14 |
| 90-99..... | 29 | 31 | 9 | 60 |
| 100-109..... | 27 | 22 | 22 | 56 |
| 110-119..... | 22 | 0 | 27 | 73 |
| 120-129..... | 15 | 0 | 13 | 87 |
| 130 or above..... | 6 | 0 | 33 | 67 |
| No. of cases..... | 107 | 21 | 20 | 66 |
| Median I. Q.'s..... | | 94 | 110 | 110 |

The only pupil testing below 80 I. Q. dropped out at the end of the first semester of 1916-1917 to go to work. All of those in group 80-89 who left school to go to work, did so by the end of the first year. Failure in school work has been recognized as the most fruitful cause of elimination from high school, but the relation between mental ability and failure in school work has not heretofore been given due consideration. At the end of two and one-half years none of those testing below 80, and only 28 percent of those testing 80-89 remain in high school. On the other hand, 100 percent of those testing 110 or over are pursuing their studies either in the Palo Alto high school or in other high schools. When the average school ratings of the different groups is taken into account the close connection between mentality and elimination will be still more apparent. The average school rating of the 21 who left school to go to work was 73 percent; of the 20 transferred to other high schools, 77 percent; and of the 66 remaining in the Palo Alto high school, 79 percent.

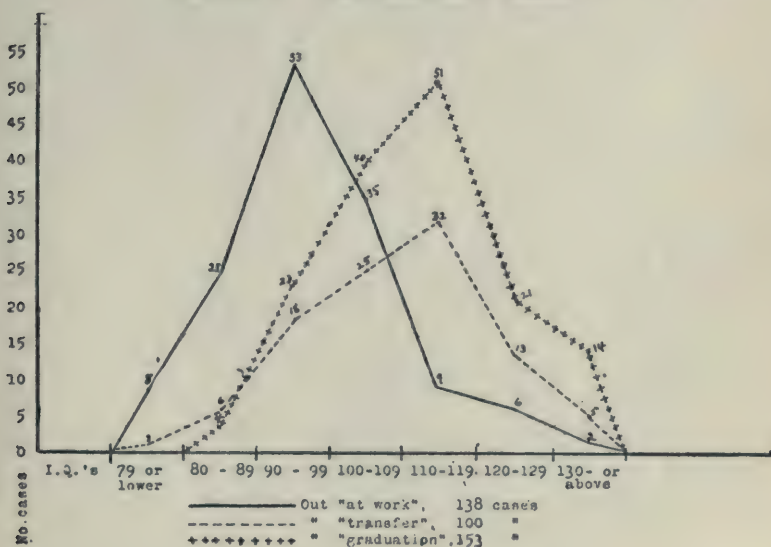
2. *Elimination among pupils tested with the Army Examination a and Alpha.*—Only five of the eight high schools where the Army Scale was applied were selected for follow-up work in connection with elimination because the records of the three Oakland high schools did not indicate whether the pupil leaving had been transferred to another high school or had gone to work. In the case of the Palo Alto, San Mateo, Redwood City, Mountain View, and Santa Clara high schools, it was comparatively easy to secure reliable data with reference to every pupil who took the test. The principals of all these schools had served in their respective positions from four to fourteen years, knew their pupils thoroughly, and had on record information concerning the movements of those who had left school since the giving of the army tests in 1917-1918.

Table VI gives for those tested with the Army Scale a distribution similar to that contained in Table V for the 107 tested with the Stanford-Binet Scale. The 955 pupils of the five high schools above mentioned have been distributed by I. Q.'s into four groups: (1) those who left high school to go to work, (2) those who transferred to some other high school, (3) those who graduated, (4) those remaining in high school on April 1, 1919. The follow-up work covered one and one-half years of school work. Figure 1 illustrates graphically the data of Table VI.

TABLE VI. DISTRIBUTION OF 955 HIGH-SCHOOL PUPILS ON
APRIL 1, 1919, WHO WERE TESTED WITH ARMY
SCALE TESTS IN 1917-1918 BY I. Q. GROUPS

| Army Scale I. Q.'s Earned 1917-1918 | Total Number of Cases | Distribution on April 1, 1919, by Percents | | | |
|---|-----------------------------|--|--|------------------------------|---|
| | | Out at Work (%) | Out, Trans- ferred to Other High School (%) | "Out by Graduation (%) | Remaining in Same High School (%) |
| 1 | 2 | 3 | 4 | 5 | 6 |
| 79 or lower.... | 13 | 61.5 | 7.7 | 0.0 | 30.8 |
| 80-89..... | 73 | 34.3 | 8.2 | 5.5 | 52.0 |
| 90-99..... | 202 | 26.2 | 8.9 | 11.4 | 53.5 |
| 100-109..... | 283 | 12.3 | 8.8 | 14.1 | 64.8 |
| 110-119..... | 221 | 4.0 | 14.5 | 23.0 | 58.5 |
| 120-129..... | 101 | 6.9 | 11.8 | 19.8 | 61.4 |
| 130 or over.... | 62 | 1.6 | 9.8 | 24.6 | 64.0 |
| No. of Cases.. | 955 | 138 | 100 | 153 | 546 |
| Median I. Q.'s..... | | 96 | 110 | 111 | 107 |
| Percent I. Q.'s below 100..... | | 62.4 | 25.0 | 17.6 | 26.5 |

FIGURE 1. ILLUSTRATING TABLE VI



In this table there appears to be a strong confirmation of the findings of Table V, as the following comparison will show:

TABLE VIA. MEDIAN I. Q.'S BY GROUPS

| | Out at Work | Out, Transferred to Other High School | Remaining in Same High School |
|---------------|-------------|--|----------------------------------|
| <i>1</i> | <i>2</i> | <i>3</i> | <i>4</i> |
| Table V..... | 94 | 110 | 110 |
| Table VI..... | 96 | 110 | 107 |

TABLE VIB. PERCENT OF PUPILS IN EACH GROUP
WHOSE I Q.'S WERE BELOW 100

| | Out at Work | Out, Transferred to other High Schools | Remaining in Same High School |
|---------------|-------------|---|----------------------------------|
| <i>1</i> | <i>2</i> | <i>3</i> | <i>4</i> |
| Table V..... | 71.4 | 20.0 | 27.0 |
| Table VI..... | 62.4 | 25.0 | 26.5 |

The agreement between the two tables is the more striking when it is remembered that the 107 high-schools pupils of Table V were all first-year pupils when the tests were given, while the 955 pupils of Table VI comprise all the classes of five different schools.

The tendency, noted in the discussion of Table II, for girls to make lower scores on the Army Scale than boys accounts for the 5.5 percent of pupils with I. Q.'s of 80-89 who appear in the "out by graduation" group. These four pupils were all girls, and their average rating in school work was 83 percent.

Table VII indicates how school marks tend to correspond to mental level as indicated by I. Q.'s earned in the Army Scale tests.

Considering the "at work" cases, it appears that only the three highest I. Q. groups were doing a passing grade of work. Lack of mental ability was perhaps the most potent cause of elimination, but it was not the only cause operating in these cases. When interest in the subjects offered, application, and ambition are lacking, high scores in the mental tests are n

TABLE VII. AVERAGE SCHOOL MARKS OF 955 HIGH-SCHOOL PUPILS, DISTRIBUTED ACCORDING TO THE CAUSES OF LEAVING

| Cause of Leaving | 79 I. Q. or Lower | 80- 89 I. Q. | 90- 99 I. Q. | 100- 109 I. Q. | 110- 119 I. Q. | 120- 129 I. Q. | 130 I. Q. or Higher | General Average |
|---------------------------------------|-------------------------|--------------------|--------------------|----------------------|----------------------|----------------------|---------------------------|--------------------|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| At work..... | 65 | 64 | 72 | 66 | 70 | 75 | 87 | 71.3 |
| Transferred..... | 59 | 67 | 76 | 72 | 82 | 85 | 86 | 75.3 |
| Graduated | | 83 | 79 | 84 | 85 | 84 | 92 | 84.5 |
| Average marks by I. Q. groups..... | 62.0 | 71.3 | 75.6 | 74.0 | 79.0 | 81.3 | 88.3 | 77.3 |

necessary guarantee of school success. Turning now to the pupils transferred to other high schools, it appears that they made better average marks under every I. Q. group, except the very lowest, than did the "at work" pupils; but it is evident that some of them transferred to other high schools because of failure in school work. Later checking up will no doubt find them out of school entirely. All of the cases in the "out by graduation" group show high average school marks, even those whose I. Q. is between 80 and 89. These are the same four cases discussed in connection with Table VI. They illustrate how necessary it is for those employing mental tests to be conservative in accepting the results of any single test as final.

When the school marks earned by all of these "out" groups are averaged a series of marks is obtained, ascending gradually from the lowest to the highest I. Q. ratings, which indicates a definite tendency for the quality of school work to correspond to the mental level indicated by the Army Scale tests.

A further following up of the 153 who graduated from the five high schools between September, 1917 and April, 1919, brought out the fact that 94 of them were continuing their education in college, university, or normal school. The median I. Q. of those going on to higher educational institutions was found to be 116. One high-school girl, who earned an I. Q. of 140 on the Army Scale, graduated from the San Mateo high school at the age of fourteen years and five months. Because of the minimum age

limit of fifteen years for entrance to the university, she was compelled to wait seven months before continuing her education. Had her case been included in the "at college" group, the median I. Q. would have been 118.

Taking into account all of the cases of high-school pupils tested either by the individual or group method of mental examination we find the following ascending scale of median I. Q.'s as a further indication that the psychological tests disclosed the *approximate* mental level of the cases discussed in the foregoing tables: first-year high-school pupils, median I. Q., 105; high-school graduates, 111; those going on to college, 116.

SUMMARY

1. Individual and group mental tests of the types described have been shown to be sufficiently reliable to justify their use as aids in determining the mental level of high-school pupils.

2. Group tests, such as Army Scale Examinations *a* and Alpha, make possible a preliminary mental survey of an entire high-school population at the beginning of a school year. The resulting raw scores will be found to be of great value in grouping the pupils according to ability. Such tests should always be supplemented with every other possible means of discovering the mental level. The results should be considered tentative and subject to revision in the light of later developments.

3. The high-school principal who makes such a preliminary mental survey of his pupils can be reasonably sure that 50 percent of those who test below normal will be eliminated within the first two years; that 25 percent additional of the subnormal group will have been transferred to other high schools because of failure in their school work; and that a negligible number will ever graduate. With this information at hand he can plan the curriculums of his pupils more intelligently. Discovering at the outset that from 15 to 30 percent of his pupils are incapable of succeeding in the conventional high-school subjects, he will undertake to make new adjustments to meet the situation. There will be fewer failures; more pupils will remain to take work that is adapted to their needs and capacities; and the high school will be less open to the charge of catering only to the intellectual aristocracy among its pupils.

RETARDATION IN ONE-ROOM RURAL SCHOOLS IN KANSAS

F. J. KELLY, *University of Kansas*

AND

A. K. LOOMIS, *Principal, Sumner County High School, Wellington, Kansas*

In 1918 Kansas still had 147,889 children enrolled in one-room rural schools out of a total of 405,319 children enrolled in all public schools. Of those in one-room schools about 72 percent were enrolled in schools having terms of seven months and almost all of the remaining 28 percent in schools having terms of eight months. The course of study in all these one-room rural schools is made out on a nine-year basis; and, if the children keep up with their classes, they are therefore still required to spend nine years to complete the course. Practically all schools in the towns and cities, on the other hand, have terms of nine months. The study here reported was one of a group of studies undertaken in the hope of reducing this inequality of opportunity.

In February, 1917 with the cooperation of the state superintendent and the county superintendents concerned, pupil's cards containing the following questions were sent to the teachers of all one-room rural schools in fourteen typical counties in Kansas.

PUPIL'S CARD

1. Are you a boy or a girl?.....
2. What is your name?.....
3. How old were you at your last birthday?.....
4. In what month were you born?.....
5. In what year were you born?.....
6. In what grade are you now?.....
7. Do you take all your studies in the same grade?.....
8. Have you ever attended any other school?.....
9. Does your father own the farm on which you live?.....
10. Do you intend to go to high school?.....
11. If you are a boy, do you intend to be a farmer?.....

A sufficient supply of the cards was sent so that each teacher had one for each pupil in her school.

Along with these cards, the following teacher's card was sent:

TEACHER'S CARD

1. County.....
 2. District No.....
 3. Teacher's name.....
 4. Teacher's address.....
 5. How many pupils are NOW on roll?.....
(Do not give enrollment from first of the year)
 6. Have you examined every pupil's card in regard to the three points mentioned in the directions for the teacher?.....
 7. How many months of school does this district have this year?.....
 8. How many months of school did this district have last year?.....
 9. How far is it from this school house to the nearest high school?.....
 10. How many years have you taught this school before this year?.....
 11. What is your monthly salary?.....
 12. State exactly what kind of certificate you hold.....
 13. How many years of teaching experience have you had before this year?.....
 14. Are you a high-school graduate.....
 15. If not a high-school graduate, how many weeks have you attended high school?.....
 16. Have you attended a state normal school?.....
 17. If so, how many weeks?.....
 18. Have you attended a college or university?.....
 19. If so, how many weeks?.....
- (OVER)

On the reverse side of the teacher's card the following instructions were printed:

TEACHER'S CARD (*Reverse Side*)

To the teachers of one-teacher rural schools:

With the approval of State Supt. W. D. Ross and Mr. J. A. Shoemaker, one of the Rural School Supervisors, I am undertaking to make a study of certain conditions in the one-teacher rural schools of Kansas. Your county superintendent is heartily in favor of the plan and I am certain that you will be willing to let the children fill out the enclosed cards at your first opportunity. The information that I have asked for on the teacher's card is very important. The teacher's card is printed on the other side of this sheet.

If you will assist in this study, please notice the following directions carefully before letting the children fill out the cards.

DIRECTIONS FOR THE TEACHERS

Let the children fill out the cards if they are old enough to do it correctly, otherwise please fill them out yourself. If you haven't received enough cards for all pupils use the cards for the older pupils, but please report the sex, name, age, and grade of every pupil for whom you send in no card.

Examine the cards after the children have filled them out to make certain of the following three points:

1. That every question has been answered.
2. That the answer to question three agrees with the answers to questions four and five.
3. That question six is answered correctly.

N. B. IN ANSWERING QUESTION SIX, COUNT A PUPIL WHO TAKES HIS WORK IN MORE THAN ONE GRADE IN THAT GRADE IN WHICH HE TAKES MOST OF HIS WORK.

Please fill out your own card carefully, answering every question. Return the cards to me in the enclosed stamped envelope within a week from the time you receive them if possible.

Very truly yours,

1110 Vermont St.

Lawrence, Kansas.

A. K. LOOMIS

(OVER)

It seems sure from these detailed instructions that the data given by the children were thus rendered reasonably accurate.

Of the fourteen counties, eleven sent in returns in time to be tabulated for this study. Of all the teachers of one-room rural schools in these eleven counties 54 percent sent in reports of all their children. The range of schools responding was from 30 percent in one county to 81 percent in another. The total number of pupil's cards received was 10,298, of which 5,267 were boys' cards and 5,031 were girls' cards. All the tabulations from these cards were made by Mr. Loomis and not by the teachers concerned.

Among the questions which the study undertakes to answer are:

1. What is the percentage of seven-month one-room schools and eight-month one-room schools by counties? (See Fig. 1.)
2. What is the amount of retardation in these one-room rural schools by counties, by ages, by grades, and by sexes? (See Figs. 2, 4, and 6.)
3. How does this retardation compare with retardation in city schools in Kansas? (See Figs. 3, 5, and 7, and Table II.)
4. How does the retardation in seven-month schools compare with the retardation in eight-month schools? (See Fig. 8.)
5. What is the cost of this retardation in terms of ambition for higher education? (See Fig. 9.)
6. What percentage of the children have spent their school lives in one school only, and what effect has the moving from one school to another had upon retardation? (See Figs. 10 and 11.)

7. What is the money cost of retardation in one-room rural schools? (See concluding paragraph.)

8. What is the money cost of the nine-year course of the one-room rural school as compared with the eight-year course of the city school? (See concluding paragraph.)

The answers to the above questions are largely contained in the tables and figures given on the following pages.

TABLE I. THE NUMBER AND PERCENT OF SCHOOLS REPORTING, AND THE TOTAL CARDS RECEIVED BY COUNTIES

| | County | | | | | | | | | | | Total |
|------------------------------------|--------|-------|-----|-------|-----|-------|-----|-----|-----|-----|-----|--------|
| | A | B | C | D | E | F | G | H | K | L | M | |
| No. of one-room rural schools..... | 78 | 125 | 91 | 98 | 91 | 113 | 89 | 68 | 72 | 106 | 108 | 1,039 |
| No. of schools reporting.... | 63 | 82 | 37 | 59 | 27 | 66 | 47 | 43 | 43 | 51 | 39 | 557 |
| Percent of schools reporting | 81 | 66 | 41 | 60 | 30 | 58 | 53 | 63 | 60 | 48 | 36 | 54 |
| No. of boys' cards rec'd. . . | 710 | 809 | 377 | 624 | 241 | 651 | 359 | 390 | 379 | 396 | 331 | 5,267 |
| No. of girls' cards rec'd. . . | 667 | 742 | 400 | 690 | 223 | 608 | 302 | 362 | 375 | 347 | 315 | 5,031 |
| Total no. of cards rec'd. . . | 1,377 | 1,551 | 777 | 1,314 | 464 | 1,259 | 661 | 752 | 754 | 743 | 646 | 10,298 |

FIGURE 1. PERCENT OF PUPILS WHO ARE IN EIGHT-MONTH SCHOOLS

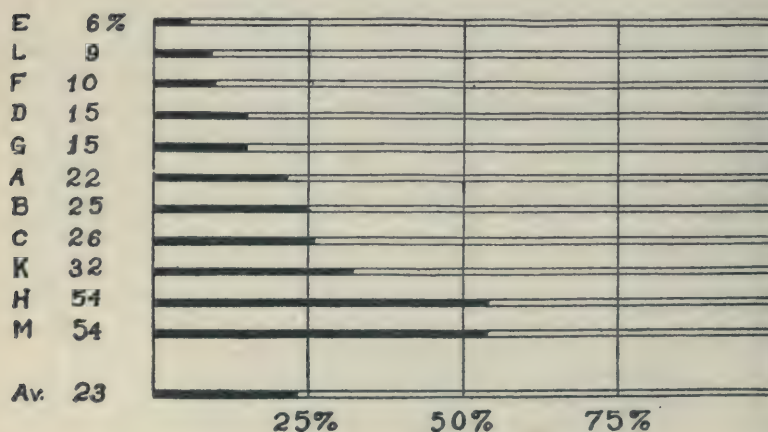


Figure 1 should be read: County E has 6 percent of its children for whom cards were received in eight-month schools, while County H has 54 percent in eight-month schools. The average of the eleven counties is also shown. Twenty-eight percent of the one-teacher rural schools of the whole state have an eight-month term. Practically all others have seven-month terms.

FIGURE 2. RETARDATION AND ACCELERATION IN 54 PERCENT OF THE ONE-TEACHER RURAL SCHOOLS IN ELEVEN COUNTIES IN KANSAS, ENROLLING 10,298 CHILDREN

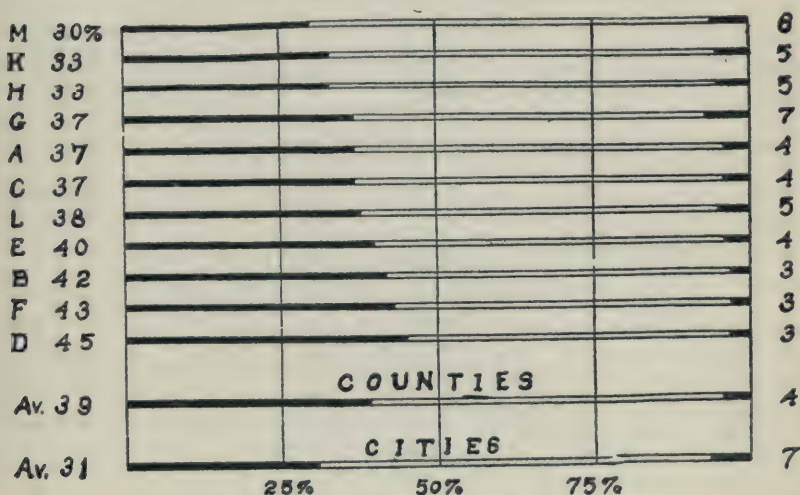


Figure 2 should be read: County M has a retardation of 30 percent and an acceleration of 6 percent. The average of the eleven counties is shown, also the average for twelve cities in Kansas in June, 1916. Normal age in grade I is six or seven at the end of school; grade II, seven or eight; grade III, eight or nine.

FIGURE 3. RETARDATION AND ACCELERATION IN TWELVE CITIES IN KANSAS, ENROLLING 28,443 CHILDREN

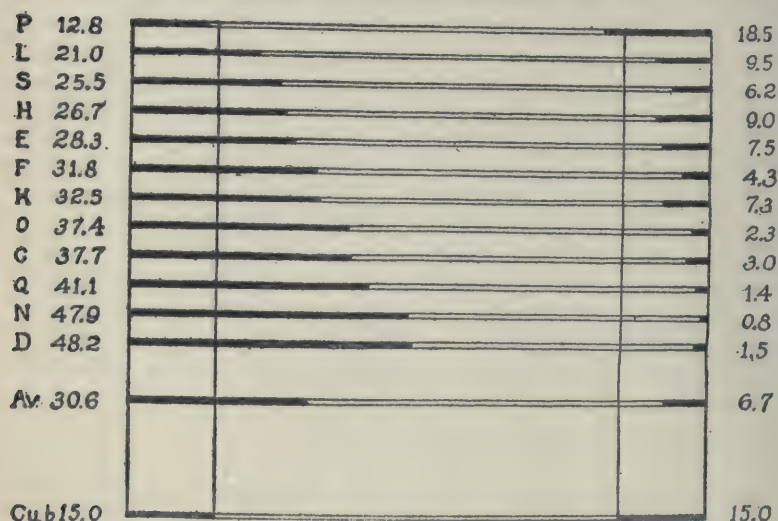


Figure 3 should be read: City P has a retardation of 12.8 percent and an acceleration of 18.5 percent. The average of the twelve cities is shown, also Cubberley's estimate of the condition under a properly adjusted course of study. The high school is included in all except N and O. Normal age in grade I is 6 or 7; grade II, 7 or 8; grade III, 8 or 9, etc.

FIGURES 4A AND 4B. RETARDATION AT EACH AGE FROM EIGHT TO FIFTEEN IN THE ONE-TEACHER RURAL SCHOOLS IN ELEVEN COUNTIES IN KANSAS

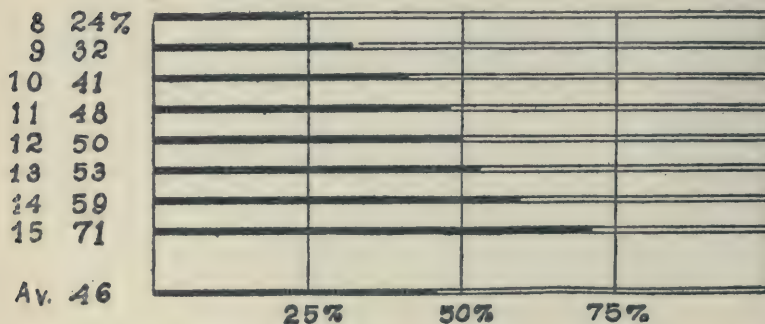


Figure 4a. Boys

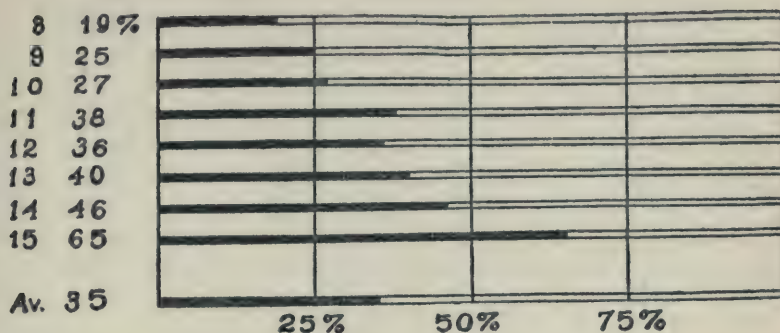


Figure 4b. Girls

Figures 4a and 4b should be read: 24 percent of eight-year-old boys are retarded and 19 percent of eight-year-old girls. The average for these ages is also shown.

FIGURES 5A AND 5B. RETARDATION AT EACH AGE FROM EIGHT
TO EIGHTEEN IN TEN CITIES IN KANSAS,
ENROLLING 26,608 CHILDREN

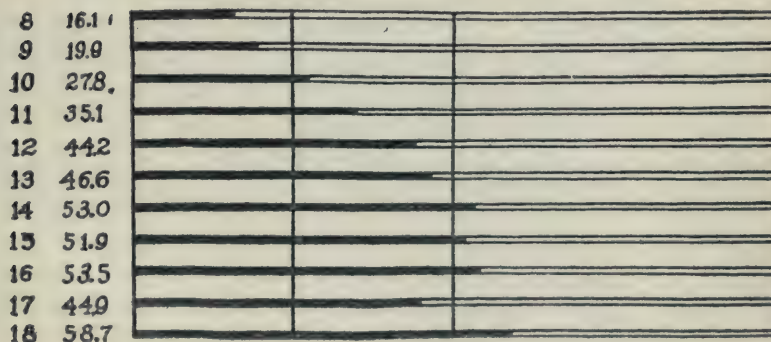


Figure 5a. Boys

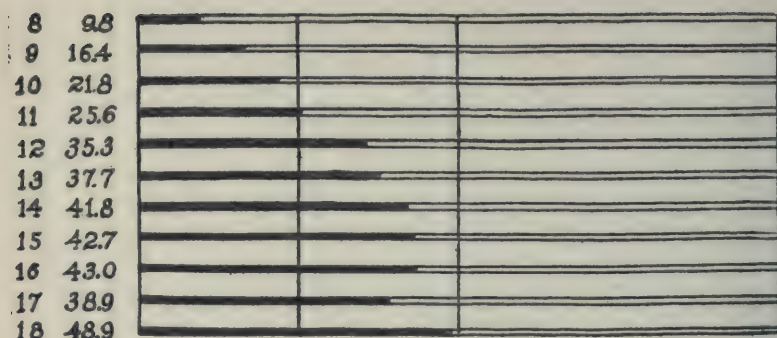


Figure 5b. Girls

Read Figures 5a and 5b as follows: 16.1 percent of eight-year-old boys are retarded, 9.8 percent of eight-year-old girls are retarded.

FIGURES 6A AND 6B. RETARDATION IN THE DIFFERENT GRADES
OF THE ONE-TEACHER RURAL SCHOOLS IN
ELEVEN COUNTIES IN KANSAS

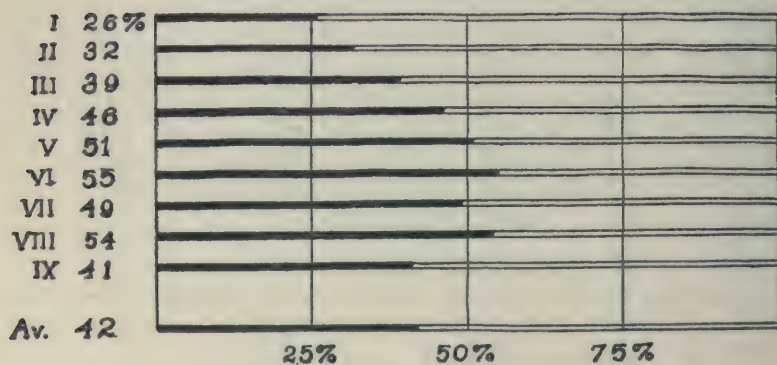


Figure 6a. Boys

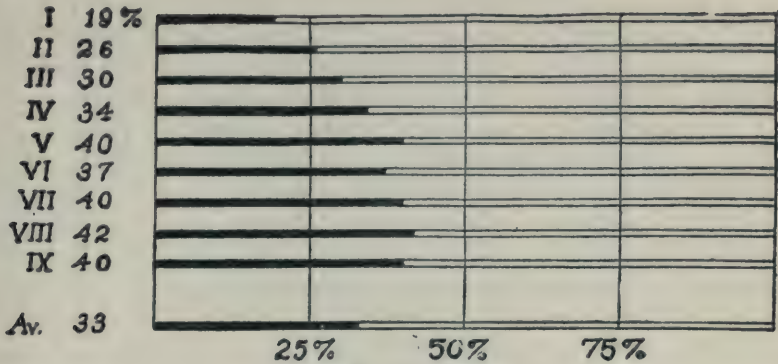


Figure 6b. Girls

Figures 6a and 6b should be read: In grade III 39 percent of the boys are retarded and 30 percent of the girls. The average for all grades is also shown.

FIGURES 7A AND 7B. RETARDATION IN THE DIFFERENT GRADES
OF THE ELEMENTARY SCHOOLS IN TWELVE
CITIES IN KANSAS

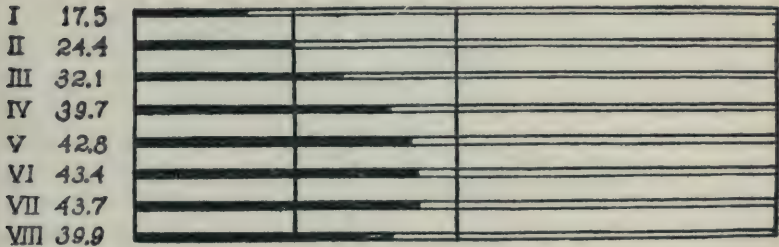


Figure 7a. Boys

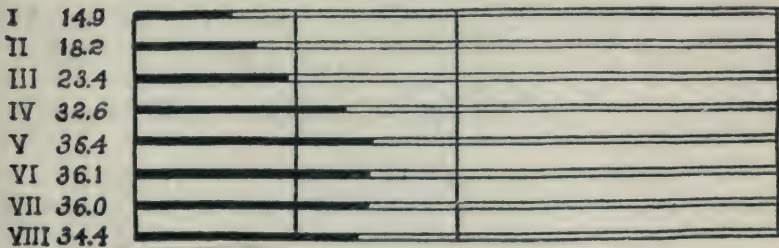


Figure 7b. Girls

Figures 7a and 7b should be read: In grade III 32.1 percent of the boys are retarded, and 23.4 percent of the girls are retarded.

TABLE II. COMPARISON OF ONE-ROOM RURAL SCHOOLS AND CITY SCHOOLS IN RESPECT TO THE PROPORTION OF PUPILS RETARDED ONE YEAR AND MORE THAN ONE YEAR

| | One-Room Rural Schools | City Schools |
|--|------------------------------|-----------------|
| Percentage retarded one year | | |
| Boys..... | 28.7 | 18.0 |
| Girls..... | 25.3 | 16.6 |
| Total..... | 27.1 | 17.7 |
| Percentage retarded more than one year | | |
| Boys..... | 17.1 | 14.1 |
| Girls..... | 10.0 | 10.6 |
| Total..... | 13.6 | 12.9 |

FIGURE 8. RETARDATION IN THE SEVEN-MONTH SCHOOLS COMPARED WITH RETARDATION IN THE EIGHT-MONTH SCHOOLS

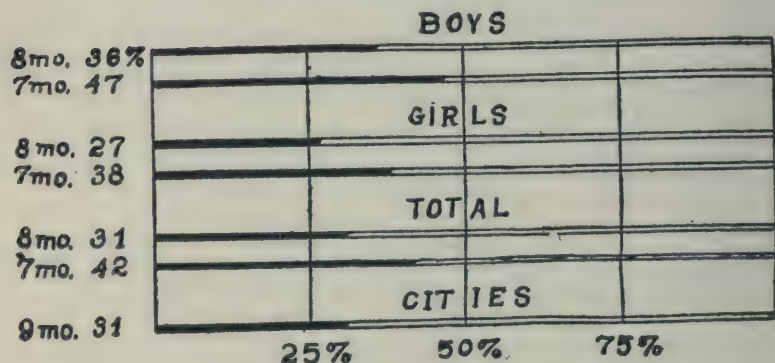


Figure 8 should be read: 36 percent of boys attending eight-month schools are retarded, while 47 percent of boys attending seven-month schools are retarded.

FIGURE 9. PERCENT OF NORMAL AND ACCELERATED CHILDREN WHO DO NOT INTEND TO GO TO HIGH SCHOOL COMPARED WITH PERCENT OF RETARDED CHILDREN WHO DO NOT INTEND TO GO TO HIGH SCHOOL

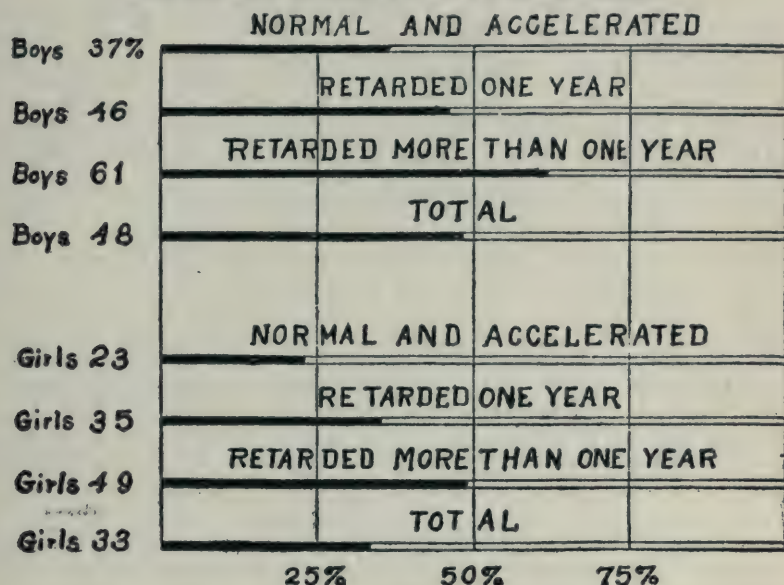
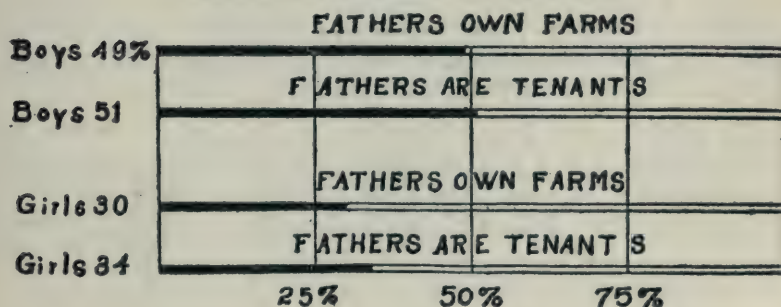


Figure 9 should be read: 37 percent of normal and accelerated boys do not intend to go to high school, 46 percent of boys who are retarded one year do not intend to go to high school, etc. Included in this comparison are 3,990 children twelve years old or older.

FIGURE 10. RETARDATION AMONG THE CHILDREN OF LANDOWNERS COMPARED WITH RETARDATION AMONG THE CHILDREN OF TENANTS



All children twelve or thirteen years of age who have attended only one school are included in this comparison. Figure 10 should be read: 49 percent of twelve and thirteen-year-old boys whose fathers own farms are retarded, and 51 percent of those whose fathers are tenants are retarded.

The comparison is based on children who have attended only one school in order to exclude from consideration at this point the effect of changing from one school to another. Twelve and thirteen-year-old children were used to avoid a fallacy resulting from using children of all ages. It is evident that the average age of all children who have attended only one school and whose fathers are tenants is less than the average age of all children who have attended only one school and whose fathers own farms. This follows from the fact that among children of tenants the proportion who have attended but one school decreases with the age of the children much more rapidly than it does among children of farm owners. Since retardation increases rapidly with age, the children of tenants fallaciously seem less retarded than the children of landowners when children of all ages are included.

FIGURE 11. RETARDATION AMONG CHILDREN WHO HAVE ATTENDED ONLY ONE SCHOOL COMPARED WITH RETARDATION AMONG CHILDREN WHO HAVE ATTENDED MORE THAN ONE SCHOOL

| ATTENDED ONLY ONE SCHOOL | | | |
|--------------------------|---------------|----------|--------|
| Boys 51% | ATTENDED MORE | THAN ONE | SCHOOL |
| Boys 61 | | | |
| ATTENDED ONLY ONE SCHOOL | | | |
| Girls 34 | ATTENDED MORE | THAN ONE | SCHOOL |
| Girls 53 | | | |
| | 25% | 50% | 75% |

All children twelve or thirteen years of age whose fathers are tenants are included in this comparison. Figure 11 should be read: 51 percent of twelve and thirteen-year-old boys who have attended only one school are retarded, and 61 percent of those who have attended more than one school are retarded.

THE MONEY COST

There remains the task of calculating the costs of retardation and of conducting seven and eight-month schools for nine years. Attention is first called to the significance of a few of the main points involved.

Let it be understood that in figuring retardation exactly the same blank was used in city schools and country schools, so that the data are entirely comparable. We wish to repeat, however, that country-school children are considered of normal age for grade, if they complete the elementary-school course before they have become sixteen, while city children are regarded as one year retarded when they complete the elementary-school course at fifteen. It is clear, therefore, that country-school children from the first grade through the ninth have progressively a larger part of a year the advantage of city children before being classified as retarded. In spite of this fact, it is noted that country children have 8 percent greater retardation as a whole than city children. Not only is that true but this difference is much greater than 8 percent among the upper ages. For example, while 51.9 percent of fifteen-year-old city-school boys are retarded, 71 percent of fifteen-year-old country-school boys are retarded, and while 42.7 percent of fifteen-year-old city-school girls are retarded, 65 percent of fifteen-year-old country-school girls are retarded. It should be noted, however, that the difference is less than 8 percent when the ninth grade of the rural school is compared with the eighth grade of the city school. This indicates that the elimination is very much greater in the country school than in the city school.

Attention is also directed to the fact that while 31 percent of all children in eight-month schools are retarded, 42 percent of all children in seven-month schools are retarded, a difference of 11 percent. This means that of the present enrollment in one-room rural schools in Kansas, practically thirteen thousand are retarded in seven-month schools who would not be retarded if they were in eight-month schools. Inasmuch as only 37 percent of boys of normal age for grade of twelve years of age or above do not intend to go to high school, whereas 46 percent of those retarded for one year and 61 percent of those retarded more than one year do not intend to go to high school, the relationship between

retardation and ambition to go to high school is direct. We recognize, of course, that many children retarded are not mentally fitted for high-school work and that retardation in their cases is not the cause of their lack of ambition. But in the case of the 11 percent who are retarded in seven-month schools, who would not be retarded if they were in eight-month schools, it may be confidently stated that retardation is the definite cause of the absence of ambition to go to high school in a considerable proportion of their cases.

Assuming that a boy or girl of thirteen or more years is worth \$150 a year as a producer, neglecting the loss as producers resulting from longer attendance of children each year between six and fourteen, and using thirteen-year-old pupils as a basis, the annual loss because of retardation in the one-room rural schools of Kansas is above \$1,800,000. One-fourth of this amount, or \$450,000, is due to the excessive retardation produced by seven-month schools above what would occur if these schools were increased to eight months. The total cost of increasing every seven-month course in Kansas to eight months would be approximately \$450,000. It is clear, therefore, that the actual saving of time in retardation alone would pay the cost of making every seven-month school into an eight-month school.

Again considering the value of a child thirteen years old and older at \$150 per year, the loss incurred in having country-school pupils required to attend nine years to complete the same course of study which would be completed in eight years (if the terms were lengthened from seven or eight months to nine months as in the city schools) can be estimated at \$2,700,000 per year. It would cost about \$1,000,000 a year to increase every seven and eight-month school to a nine-month school. There would be a saving of \$1,700,000 on this score.

Naturally a very much larger proportion of the children would complete the elementary-school course if it were only eight years in length rather than nine, thus securing very much larger advantages of education than are now possible to country children. Kansas is therefore paying very dearly indeed in money, in educational opportunity, and in ambition, by persisting in maintaining country schools of seven and eight-month terms.

A GROUP SCALE OF INTELLIGENCE FOR USE IN THE FIRST THREE GRADES: ITS VALIDITY AND RELIABILITY

LUELLA WINIFRED PRESSEY
Indiana University

I. THE STATISTICAL PROBLEMS PRESENTED BY THE SCALE

In a recent number¹ of the *Journal of Educational Psychology*, the writer has presented a general descriptive account of the purpose, nature, and development of the "Primer Scale." At the time of writing that paper, little statistical evidence had been gathered as to the "validity" and "reliability" of the scale. By "validity" the writer means the extent to which the scale measures what it is supposed to measure—i.e., general intelligence; by "reliability" is meant the consistency or accuracy of the measures obtained. Thus, a scale might give very consistent measures that were not valid—that did not measure the qualities the scale was designed to measure—or, a scale might give reasonably valid measures of some quality, but the measures might be so inconsistent as to make the scale largely unsatisfactory. In other words, these two characteristics of a scale might exist in varying combinations with each other and should be kept separate in considering the value of the results. It is the purpose of the present paper to present the evidence that has been found, thus far, bearing on these two important points of (1) the validity and (2) the reliability of the Primer Scale.

II. THE VALIDITY OF THE PRIMER SCALE

1. *Sources of data.*—For a general description of the four tests composing the scale, the reader is referred to the above mentioned article. It may be briefly said that the first test is one of form discrimination in which dots are arranged in patterns; the second is a test of similarities and differences, using pictures of familiar objects as materials; the third is a form board test, arranged so that it can be presented on paper; and the fourth is a test of absurdities, again using pictures.

¹ Pressey, L. W. "A group scale of intelligence for use in the first three grades," *Journal of Educational Psychology*, 10:297-308, September, 1919.

The tests have been given to about twenty-five hundred public school children (June, 1919) in the first three grades;² also to 64 "children" with a mental age on the Stanford-Binet Scale of between six and eight in the state school for the feeble-minded in Minnesota. Many of the school children tested had been previously "Bineted."

2. *Methods*.—The writer has used three methods for arriving at a conclusion concerning the validity of the scale. (1) One hundred and twenty children who had been "Bineted" were divided into five groups on the basis of their Intelligence Quotients (I. Q.'s) as determined by their scores in the Binet Scale (Stanford Revision). The scores of these groups in the Primer Test were then examined for consistency with the groupings according to the Binet ratings. (2) Three correlations with the Binet ratings were calculated: first, the correlation between the Binet "mental age" and the Primer Test score for the 64 institutional feeble-minded cases; second, a similar correlation for 148 children in the primary grades of Council Bluffs, Iowa; and third, a correlation for 57 unselected six-year-old children in the first grade at Washington, Indiana. (3) The writer chose this last group for a more detailed analysis by the method of partial correlation. The separate tests were intercorrelated, each test was correlated with the Binet ratings, and the partial coefficients and regression equation were calculated.

3. *Results*.—The results of the first method are indicated in Table I which gives the distribution of scores on the Primer Scale of 120 unselected six-year-old children, for whom the writer had Binet ratings. The cases have been divided into five groups on the basis of the Binet I. Q.'s. Group I (the first column) includes the scores of those children with an I. Q. of 125 or above; Group II contains those children who obtain an I. Q. of 110-124; Group III those with an I. Q. from 90-109; Group IV those from 76-89; and Group V those with an I. Q. below 76. The five groups might be termed, very superior, superior, average, inferior, very inferior.

There is some over lapping from one group to another, but the only distribution for which the amount of over lapping is surprising is that of the "average" group (Group III) whose distribu-

² The writer wishes to express her obligations to the school superintendents of Washington, Bedford, and Bloomington, Indiana, and of Council Bluffs, Iowa; also, to Dr. F. Kuhlmann, Psychologist of the School for the Feeble-minded at Faribault, Minnesota for the materials here presented.

tion over laps that of every other group. The two extreme groups (I and V) are quite distinct from each other—that is, the lowest score in Group I is above the highest score in Group V. With the exception of one case, Groups I and II combined (i.e., the really bright children according to the Binet Scale) show no over lapping of test scores with the Groups IV and V combined (i.e., with the really dull children). The horizontal lines in Table I show the location of the 75-percentile, the median and the 25-percentile for the total distribution of these unselected children—i.e., the standard norms derived from the complete surveys made. As will be seen, the children in Group I make no scores below the 75-percentile for their age, while the children in Group V, with one exception, make no scores above the 25-percentile.

The lowest group is the most compact of the five. There is, in general, not as much over lapping of one group upon the other at the lower end of the distribution as at the upper end. This is understandable, as the brilliant children might not give their best performance and would thus make a lower score than they really should; while, on the other hand, it is not to be expected that adventitious circumstances would cause a dull child to be rated too high to the same degree. On the whole, the Primer Scale seems to have marked off degrees of ability as well as could be expected of a group test given to such young children. In this connection, it might be emphasized that some six-year-old children who make a low score on a group test receive a higher rating on the Binet examination because of the very different testing environment. This aids in explaining the dropping down of scores in the first three groups into the distributions of the lower groups. Some of the children making these scores lack, not intelligence, but independence. When asked to perform certain tasks in a group, they do not have the initiative to do their best; whereas they show a Binet rating more nearly commensurate with their true intelligence, because the examiner can constantly recall their attention and their efforts to the task in hand. The writer feels that the relationship shown in Table I is probably as high as can be expected between a group test and an individual test,³ especially with such young children.

³ Lowell, Frances. "A group intelligence scale for primary grades," *Journal of Applied Psychology*, 3:215-48, September, 1919. The scale described in this article is a group form of the younger years of the Binet; yet the correlation between group score and the rating on the Binet Scale when given to the same children individually is only +0.75.

TABLE I. THE DISTRIBUTION OF SCORES MADE BY 120 SIX-YEAR-OLD CHILDREN ON THE PRIMER SCALE, GROUPED ACCORDING TO I. Q. ON THE STANFORD-BINET SCALE

| Primer Scale Scores | Group I (I. Q. above 125) | Group II (I. Q. 110-124) | Group III (I. Q. 90-109) | Group IV (I. Q. 76-89) | Group V (I. Q. below 76) |
|----------------------|------------------------------------|--------------------------------|--------------------------------|------------------------------|-----------------------------------|
| | 1 | 2 | 3 | 4 | 5 |
| 80 | 1 | | | | |
| 78 | | | | | |
| 76 | | | | | |
| 74 | | | | | |
| 72 | 2 | | 1 | | |
| 70 | 1 | | | | |
| 68 | | | | | |
| 66 | 1 | 2 | | | |
| 64 | 1 | 1 | 1 | | |
| 62 | 4 | 4 | 2 | | |
| 60 | 3 | 2 | 1 | | |
| 58 | 2 | 2 | 3 | | |
| 56 | 4 | 1 | 4 | | |
| 54 | | 1 | 4 | | |
| 52 | | | 5 | | |
| 50 | | 1 | 1 | | |
| 48 | 2 | 2 | 1 | | |
| <i>75-percentile</i> | | | | | |
| 46 | | 3 | 3 | 1 | |
| 44 | | | 5 | | |
| 42 | | 2 | 4 | | |
| 40 | | 1 | 2 | | |
| 38 | | 1 | 3 | | |
| <i>Median</i> | | | | | |
| 36 | | 2 | 1 | | |
| 34 | | | 1 | | |
| 32 | | | 1 | | |
| 30 | | | 2 | | |
| 28 | | | 4 | | 1 |
| 26 | | | | | |
| 24 | | | 2 | | |
| 22 | | | | | |
| 20 | | | 1 | | |
| 18 | | | | | |
| <i>25-percentile</i> | | | | | |
| 16 | | | 1 | | 1 |
| 14 | | | 2 | 2 | |
| 12 | | | | | |
| 10 | | | 1 | 2 | |
| 8 | | | 1 | 2 | |
| 6 | | | 1 | 1 | 1 |
| 4 | | | | | 1 |
| 2 | | | | | 1 |
| 0 | | | | | 3 |
| No. cases... | 21 | 25 | 58 | 8 | 8 |
| Medians | 61 | 55 | 45 | 11 | 4 |

The second method of studying the validity of the Primer Scale involved the calculation of the Pearson coefficient of correlation between scores derived from it and scores obtained by using the Binet Scale. The crude coefficients of correlation for each of the three groups of children are as follows:

| | |
|---|-------|
| 64 children from a feeble-minded school | +0.75 |
| 148 children in primary grades | +0.66 |
| 57 unselected children 6 years of age | +0.62 |

In all cases, the correlation is between group test scores and "mental ages" on the Binet examination.⁴ The correlations are, the writer feels, as high as can be expected from the very different nature of the testing conditions for the two types of tests (group and individual).

The scores of the 57 unselected six-year-old children were subjected to further analysis by means of partial correlation. The crude coefficients of the scores of each of the four tests which make up the Primer Scale with each other and with the Binet ratings are given in Table II.

TABLE II. CRUDE COEFFICIENTS OF CORRELATION BETWEEN TESTS OF THE PRIMER SCALE AND THE BINET SCALE, TAKEN PAIR BY PAIR

| Tests | 1 | 2 | 3 | 4 | LEGEND |
|-------|------|------|------|------|--------------------------|
| 1 | | | | | 1 = Test 1, Primer Scale |
| 2 | 0.59 | | | | 2 = Test 2, " " |
| 3 | 0.48 | 0.46 | | | 3 = Test 3, " " |
| 4 | 0.57 | 0.55 | 0.33 | | 4 = Test 4, " " |
| 5 | 0.53 | 0.45 | 0.42 | 0.47 | 5 = Binet Scale |

This table should be read: the correlation between Tests 1 and 2 of the Primer Scale was found to be 0.59; between Tests 1 and 3, 0.48; etc.

It will be noticed that the coefficients of correlation between the tests of the Primer Scale are rather high, except between the third and fourth. Because of this difference in the relationship

⁴ The I. Q., which is frequently used for such correlations, is a percent statement. No similar statement was readily possible with the Primer Scale. Therefore, it was necessary to use *total* scores on both scales. Even in correlating the scores of children of the same age, the error introduced in using I. Q. is considerable, as the basis for calculating the I. Q. may range (in the case above) from 6 years and 0 months to 6 years and 11 months—or nearly a whole year.

between the tests, it is especially desirable to find the partial correlations of each test with the Binet Scale, to determine just what the independent value of each test might be. The resulting partial coefficients of correlation are as follows:

$$r_{51.234} = 0.2436$$

$$r_{52.134} = 0.0951$$

$$r_{53.124} = 0.1834$$

$$r_{54.123} = 0.2006$$

These partial coefficients should be read: the correlation of 5 (Binet rating) with 1 (Test 1), the other three tests being constant, is 0.2436; of Binet rating and 2 (Test 2), the other three tests being constant, is 0.0951.

The partial coefficients are interesting in themselves aside from the part they play in calculating the regression equation.⁵ The relative standing of the four tests in their relation to the Binet Scale, has not changed materially from that indicated by the original correlations (Table II). The first test still has the highest correlation, and the order of decreasing coefficients is now the fourth, the third, and the second according to the partial correlations shown above. The second test seems to contribute almost nothing to the correlation of the whole scale with the Binet Scale, which is not already contributed by some other test. When the other tests are held constant, its correlation with the Binet Scale becomes practically zero (0.0951). Thus, the second test is, according to this criterion, the least valuable test of the scale⁶ and the first test is the most valuable.

III. THE RELIABILITY OF THE PRIMER SCALE

1. *Materials*.—The papers of all the children in the first and second grades, and of all the six- and seven-year-old children in

⁵ The regression equation of the Binet ratings on the four tests of the scale was found to be the following:

$$x_5 = 0.5719x_1 + 0.2142x_2 + 0.2844x_3 + 0.5380x_4,$$

where x is the deviation of a pupil's score in Test 1 from the average, x_2 his deviation from the average in Test 2, etc. x_5 is the most probable deviation from the average Binet score. The scores of the original 57 papers used in the derivation of the equation were weighted in accordance with the above equation. This weighting raised the correlation from 0.62 to 0.65.

⁶ Since writing the above, the writer has tried a new method of scoring the second test, which takes account of possible chance successes, and has obtained a correlation of 0.52 instead of 0.45 as given in Table II. This difference in correlation might make some difference in the relative standing of the second test as revealed by partial correlation.

the city most recently tested were chosen as the material from which to estimate the reliability of the scale.

2. *Methods*.—The degree of the reliability of the measures obtained by the scale and of its separate tests has been calculated in two ways. Coefficients of reliability have been found for each of the four groups just mentioned (grades I and II and ages six and seven). As there was only one form of the scale, the coefficients of reliability were calculated by dividing the scale into halves (using alternate items), correlating the scores of the half tests, and applying Brown's formula.⁷ This formula represents the extent to which the amalgamated results of two tests would correlate with a similar amalgamated series of two other applications of the same test. The formula is:

$$r_2 = \frac{2r_1}{1+r_1}$$

Both the total scale and each test have been treated in this way for the four age and grade groups. (2) The reliability has been calculated in terms of the probable error (P. E.), by the following method. Each pupil's score on one half of the scale was subtracted from his score on the other half. These differences were then averaged, giving an "average difference."

$$\begin{aligned} \text{The P. E. of a scale} &= \frac{0.8453}{\sqrt{2}} \times \text{Av. Diff.} \\ &= 0.5978 \times \text{Av. Diff.}^8 \end{aligned}$$

This formula applies whether the two tests from which the average difference is found are two separate equivalent tests or two halves of the same test. If two separate forms of the same test had been used, the above formula (P. E. = $0.5978 \times \text{Av. Diff.}$) would give the correct value, but in the case of the two halves of the same test, a further calculation must be made. The formula gives the P. E. of the half tests, but what is desired to obtain finally, is the P. E. of the *whole* test—not the P. E. of one of its halves. The P. E. of a test increases as the square root of its length; that is, in this case, the whole test is *twice* as long as its halves, and its P. E. consequently increases by $\sqrt{2}$. The P. E. of the whole test would, then, be equal to that of the

⁷ Brown, William. *Essentials of mental measurement*. New York: G. P. Putnam's Sons, 1911, p. 101.

⁸ Thorndike, E. L. *An introduction to the theory of mental and social measurements*. New York: Teachers College, Columbia University, 1916, pp. 190-93.

half test, multiplied by the $\sqrt{2}$. The resulting formula for deriving the P. E. of a test from that of its halves becomes:

$$\text{P. E.} = 0.5978 \times \text{Av. Diff.} \times \sqrt{2}^9$$

3. *Results.*—The coefficients of reliability for grades I and II and for ages six and seven are given in Table III. For grade I and ages six, the first column represents the reliability of the scale, including all zero scores; the second column gives the reliability

TABLE III. THE COEFFICIENTS OF RELIABILITY BETWEEN THE TWO HALVES OF THE PRIMER SCALE FOR GRADES I AND II AND FOR AGES SIX AND SEVEN, AFTER APPLYING THE FORMULA:

$$r_2 = \frac{2r_1}{1+r_1}$$

(In this formula r_1 is the coefficient of correlation obtained by correlating the two halves of the scale and r_2 is the coefficient of reliability given in the table below.)

| | GRADE I (142 CASES) | | AGE SIX (90 CASES) | | Grade II (120 Cases) | Age Seven (103 Cases) |
|-------------------|----------------------------|---------------------------|----------------------------|---------------------------|----------------------------|-----------------------------|
| | Zero Scores Included | Zero Scores Omitted | Zero Scores Included | Zero Scores Omitted | | |
| | 1 | 2 | 3 | 4 | 5 | 6 |
| Total score | 0.96 | 0.95 | 0.95 | 0.94 | 0.89 | 0.92 |
| Test I | 0.94 | 0.90 | 0.92 | 0.89 | 0.89 | 0.84 |
| Test II | 0.92 | 0.81 | 0.88 | 0.81 | 0.67 | 0.68 |
| Test III | 0.89 | 0.81 | 0.88 | 0.81 | 0.93 | 0.93 |
| Test IV | 0.82 | 0.78 | 0.84 | 0.79 | 0.54 | 0.75 |

omitting the zero scores. Neither method is satisfactory. If the test were to be given again, some but probably not all, of those children making zero scores would make a higher score. Since some of them would, the first column of coefficients, including all the zero scores, is too high: but, since not all of them would, the second column of coefficients, omitting all zero scores, is too low. Thus the true coefficient probably lies somewhere between the two.

The coefficients of reliability are uniformly high for the whole scale and for the first and third tests, but vary considerably for

⁹ The writer is indebted to Dr. Walter S. Monroe for this method of determining the P. E. of a scale from the P. E. of its halves; as well as for many other helpful suggestions and criticisms in the course of the statistical analysis.

the second and fourth tests in the different groups. The reliability, as measured by the correlation of one half of a scale to the other, is in part a measure of the consistency of the performance of the pupils, and in part a measure of the consistency of the test from item to item. It is the judgment of the writer that the reliability coefficient is mainly an index of the latter factor. If the halves of a test correlate highly, it is evident that the items composing the two halves demand the same, or approximately the same kind of mental effort. For Tests I and III this consistency is high; these tests are both "form" tests, they demand no information of any kind, and it apparently makes little difference which set of items is used. Tests II and IV call for information of a practical sort; the two halves of these tests correlate rather poorly—sometimes very poorly. Evidently, practical information is not at all the unitary thing that form discrimination is; and the kind of information demanded by one item is different from that demanded by another. If still another form of the test were to be used, it is probable that these two tests would still correlate poorly, since these new items would involve still different informational elements.

The P. E. of the scale was derived in the manner explained above. The P. E. of grade I (142 cases) was 2.39 "points," or units; that of age six (90 cases) was 2.20 "points," or units on the scale. Since the difference between the median of one age and the median of the next age is, on the average, twelve "points," or units, a P. E. of 2.39 points would equal 2.39 months, and a P. E. of 2.20 points would equal 2.20 months. Again, since there are about twelve points from one age to the next, a statement, for example, that a given child scores at the median for the age above his own, does not seem to be invalidated by the unreliability of the scale, since the differences between the successive medians are well beyond the P. E. of the scores.

The P. E. of the Stanford-Binet Scale has been estimated by Otis to be about three and a half months.¹⁰ The P. E. of the Primer Scale is even less. If the same reliability can be obtained with a group scale as with an individual scale, the group scale can,

¹⁰ Otis, A. S. "An absolute point scale for the group measurement of intelligence," *Journal of Educational Psychology*, 9:341, June, 1918.

to some extent, take the place of the individual scale, at least for the preliminary sorting of pupils.

On the whole, the scale seems very reliable—unusually so when the ages of the children are considered. Considerable confidence may, therefore, be placed in the accuracy of the scores obtained.

IV. SUMMARY

1. The writer presents data to show the reliability and the validity of the Primer Scale.

2. The validity is shown (1) by the distribution of scores made on the Primer Scale by children who have been classified according to I. Q. on the Binet examination; (2) by the correlations for various groups of children between the group scale and the Binet ratings; and (3) by partial coefficients of correlation.

3. The reliability is shown by (1) the coefficients of reliability for different ages and grades and (2) by the P. E. of the whole scale, as derived from the P. E. of the half scale.

4. It is concluded that the Primer Scale is reasonably valid and reliable.

THE APPLICATION TO TABLES OF DISTRIBUTION OF A SHORTER METHOD FOR COMPUTING COEFFICIENTS OF CORRELATION

LEONARD P. AYRES
Russell Sage Foundation

In the March number of the JOURNAL OF EDUCATIONAL RESEARCH there appeared an article by the writer on a shorter method for computing the coefficient of correlation. The method described avoids the necessity for finding the deviations of the separate measures from their averages and so eliminates the necessity for watching for unlike signs and making the series of corrections to allow for differences between the guessed averages and the true averages of the two series. The present article explains the application of the method to tables of distribution.

The new method is based on the principle that in any series of numbers the sum of the squares of the deviations from the average is equal to the sum of the squares of the numbers in the series, minus the product of the total of the series and its average. The formula for this method is as follows:

$$r = \frac{\begin{array}{c} \text{Sum of products of subject and relative items} \\ \text{minus} \\ \text{av. of subject items} \times \text{tot. of relative items} \end{array}}{\sqrt{\begin{array}{c} \text{Sum of squares of sub. items} \\ \text{minus} \\ \text{av.} \times \text{total of these items} \end{array} \times \begin{array}{c} \text{Sum of squares of rel. items} \\ \text{minus} \\ \text{av.} \times \text{total of these items} \end{array}}}$$

In using this method the multiplications are made first and the corrections are made separately. For two simple series of paired measures the operations may be illustrated as follows:

| | <i>Sub.</i> | <i>Rel.</i> | <i>Sub.</i> ² | <i>Rel.</i> ² | <i>Sub.</i> × <i>Rel.</i> |
|------|-------------|-------------|--------------------------|--------------------------|---------------------------|
| | <hr/> | <hr/> | <hr/> | <hr/> | <hr/> |
| | 2 | 3 | 4 | 9 | 6 |
| | 3 | 2 | 9 | 4 | 6 |
| | 4 | 4 | 16 | 16 | 16 |
| | 7 | 4 | 49 | 16 | 28 |
| | 9 | 7 | 81 | 49 | 63 |
| | <hr/> | <hr/> | <hr/> | <hr/> | <hr/> |
| Tot. | 25 | 20 | 159 | 94 | 119 |
| Av. | 5 | 4 | | | |

$$\begin{aligned}
 5 \times 20 &= 100 \\
 5 \times 25 &= 125 \\
 4 \times 20 &= 80 \\
 119 - 100 &= 19 \\
 159 - 125 &= 34 \\
 94 - 80 &= 14
 \end{aligned}$$

$$r = \frac{19}{\sqrt{34 \times 14}} = \frac{19}{\sqrt{476}} = \frac{19}{21.82} = 0.87$$

The procedure is similar when the method is applied to a distribution table. To illustrate it the case may be taken of a supposititious age and grade table for the elementary grades of a school system in which one hundred children enter the first grade each year at the age of six, are promoted by annual promotions of 80 percent so that eight in each ten go forward and two repeat, and none drop out of school during the elementary years. The problem is to compute the coefficient of correlation between the ages and the grades of the children. The work is as shown in the accompanying computations. The letters "S" and "R" stand for "subject" and "relative."

| S | R-Ages | | | | | | | | T | ST | SST |
|--------|--------|-------|-------|-------|--------|--------|--------|--------|--------|--|--------|
| | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | | | |
| Grades | | | | | | | | | | | |
| 1 | 100 | 20 | 4 | 1 | | | | | 125 | 125 | 125 |
| 2 | | 80 | 32 | 9 | 2 | | | | 123 | 246 | 492 |
| 3 | | | 64 | 39 | 15 | 5 | 1 | | 124 | 372 | 1,116 |
| 4 | | | | 51 | 42 | 20 | 8 | 2 | 123 | 492 | 1,968 |
| 5 | | | | | 41 | 42 | 24 | 11 | 118 | 590 | 2,950 |
| 6 | | | | | | 33 | 41 | 27 | 101 | 606 | 3,636 |
| 7 | | | | | | | 26 | 39 | 65 | 455 | 3,185 |
| 8 | | | | | | | | 21 | 21 | 168 | 1,344 |
| T | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 800 | 3,054 | 14,816 |
| RT | 600 | 700 | 800 | 900 | 1,000 | 1,100 | 1,200 | 1,300 | 7,600 | Av. of sub. items $\frac{3,054}{800} = 3.818$ Av. | |
| RRT | 3,600 | 4,900 | 6,400 | 8,100 | 10,000 | 12,100 | 14,400 | 16,900 | 76,400 | | |
| ΣSf | 100 | 180 | 260 | 340 | 422 | 503 | 583 | 666 | | Av. of rel. items $\frac{7,600}{800} = 9.5$ Av. | |
| R(ΣSf) | 600 | 1,260 | 2,080 | 3,060 | 4,220 | 5,533 | 6,996 | 8,658 | 32,407 | | |

$$3,054 \times 9.5 = 29,013$$

$$3,054 \times 3.818 = 11,660$$

$$7,600 \times 9.5 = 72,200$$

$$32,407 - 29,013 = 3,394$$

$$14,816 - 11,660 = 3,156$$

$$76,400 - 72,200 = 4,200$$

$$r = \frac{3,394}{\sqrt{3,156 \times 4,200}} = \frac{3,394}{\sqrt{13,255,200}} = \frac{3,394}{3,641} = 0.932$$

In the illustrative example the columns headed "ST" and "SST" give the items of the subject series and the squares of these items. They are found by the simple process of multiplying the numbers in the total column by those in the column headed "S" at the extreme left of the table and then multiplying these new totals by the "S" numbers once more. A similar process gives the "RT" and "RRT" rows at the bottom of the table.

The next step is to find the products of the paired values. This is done in the two last rows at the bottom of the table. The first is headed " ΣSf " and gives for each column the sum of the products of the numbers in the "S" column and their corresponding frequencies in the body of the table. Thus the sum of the products for the second column is 180 and consists of 1×20 plus 2×80 .

After all these sums of products are found and entered in the row headed " ΣSf ," they are multiplied by the "R" numbers at the heads of the columns and the results entered in the lowest row headed $R(\Sigma Sf)$.

The averages are found by dividing the total of the subject items (3,054) by the number of cases (800) which gives 3.818. Similarly the sum of the relative items (7,600) is divided by 800 to give the other average of 9.5.

The next step is to find the three "total \times average" products and to make the three corrections which appear after the table. The first product is that of the total of one series (3,054) and the average of the other (9.5), or 29,013. This is subtracted from the sum of the products of the paired values (32,407) giving the remainder 3,394. This equals the sum of the x times y products as found in the conventional formula:

$$r = \frac{\Sigma xy}{\sqrt{\Sigma x^2} \sqrt{\Sigma y^2}}$$

Similarly the sum of the squares of the items of the subject series (14,816) is corrected by subtracting from it the total of the series times its average. The corresponding operation is done for the relative series. The final operations follow the formula given at the beginning of the article and are the same as those of the conventional procedure.

The work is greatly facilitated by the use of an adding machine. The writer did all the work of the example given in twenty minutes after the age-grade table had been written down and checked it by repeating all the operations in fifteen minutes. This was done with the aid of a comptometer—a small, non-listing, adding machine which can also be used for multiplying.

On work such as that given in this article the new method is found to be in practice much faster than the old one and errors in computation are very much less likely to be made.

Erratum

In Dr. Ayres' first article in the March number of the JOURNAL OF EDUCATIONAL RESEARCH at page 220, middle of page, read:

"Second correction = $214 - (6 \times 30) = 214 - 180 = 34$."

Editorials

THE ACHIEVEMENTS OF THE SCHOOL YEAR

As the school year draws to a close, do the plans of superintendents and teachers involve taking an inventory of the achievements of the year? The merchant, the manufacturer, or the banker (and we are beginning to include the farmer) who does not at least once a year take an inventory of his business is generally considered to be headed toward certain failure. If when dealing with such tangible things as clothing, drygoods, machinery, automobiles, corn, cattle, and hogs, it is necessary to measure accurately the results of one's activities during the year, how much more is it necessary to render an accounting when one is dealing with the less tangible achievements of the school room.

In the records of the school census, attendance, enrollment, withdrawals, and promotion the superintendent has at hand much of the crude data which an inventory of his business requires. Many superintendents have given, at times during the year, one or more standardized tests. The results from these furnish additional crude data. However, in order to secure a satisfactory inventory, there should be an appraisal of the achievements of the pupils during the school year. The way to do this is to plan a comprehensive testing program for use toward the close of the school year. With the large array of standardized tests now available, the superintendent can secure the instruments needed for measuring many of the important achievements accomplished during the year.

The labor involved in the preparation of an inventory of the year's work will perhaps appall many schoolmen. Some will doubtless claim that they have been successful in the past without rendering an account of the year's work. In reply to these it is only necessary to say that there was a time when many merchants were successful who never made an inventory of their business.

But the time has now passed in the business world when a man can expect to be successful without it. If the time has not arrived in educational activities when such a procedure is necessary for success, the time is rapidly approaching. It, therefore, behooves those superintendents who desire to be rated among the successful ones to plan definitely for the rendering of an accounting of the year's work.

W. S. M.

THE GIFTED CHILD

For years serious attention has been given to the handling of backward and defective children. This involves methods of instruction, curriculum making, promotion, and segregation. Such efforts are worth while, and we do not wish to minimize their importance. We do wish, however, to insist that we must give more attention than has heretofore been given to the education of children of superior mental ability.

Superior children who are merely carried along through the grades under the customary "time service" spend many days of their school life during which their minds are being injured rather than educated. Many a boy, instead of letting his brain power run idle in the classroom while slower minds labor to understand problems that present no difficulty to him, would be better off playing in the field. Ought a bright child be penalized because he is bright? We speak of equal rights; but equal rights involve equal opportunities. And such children are robbed of their opportunity for self-realization.

Conservation of power and reduction of waste is an important social ideal. A lot of the brain power in our schools is daily going to waste because it is idling far below its normal working capacity. Genius which spends itself on inferior work is wasted. Our educational practice needs a severe earthquake that will shake us loose to the recognition of this waste.

Recently we have heard more about "doing something for superior children," but actual efforts to discover such children have been comparatively few and spasmodic. With the present development of mental testing we have an effective tool by which teachers and school administrators may make systematic search for these children. If such an instrument is not used, each gifted

child must force himself upon the attention of someone who has backbone enough to break away from tradition. Classroom procedure is not organized so as to favor a child's forcing himself into recognition. Moreover, in many places even if superior ability is perfectly evident, nothing short of a miracle will cause the traditional system to break its regularity of habit.

The latest development in mental testing is the group test. It can be given to hundreds of children in a single day. It affords impersonal data from which to estimate a child's capacity, and it thus removes the decision from the field where the personal feelings of teachers or of fond parents play a part. Experience proves that these group tests, in spite of the ease with which they may be administered, have a high degree of reliability. The results can be used for "spotting" individuals; and wherever there are serious differences between test results and other data, the child may be given further study or additional tests. These group tests ought to come into immediate and extensive use for the discovery of superior mental capacity in all classes from grade three up. Indeed, some group tests promise to be fairly reliable even below the third grade. Individual mental tests should be used on all school children who are too young to be tested by means of a group test.

No superintendent or principal should be satisfied with anything less than a complete and systematic search of his school for superior ability. Having found such ability, he should devise ways and means to handle the administrative problems of education so that this ability may be put to work. This is not altogether easy. It means that more thinking and more effort must be put into both administration and teaching. The problem can be solved, and it is being solved in many places. Various plans are being used but none of them is as important as the clear realization that superior mental capacity should be used and trained instead of being allowed to waste away in the formation of habits of idleness and mischief.

With such a realization, intelligence properly applied will meet the problem. Whether we like the work or not, it must be done. Let us face the task squarely. With due apology to advertising, we may say "do it now, for eventually you must."

VIRGIL E. DICKSON

Reviews and Abstracts

COURTIS, STUART A. *The Gary public schools: measurement of classroom products.* New York: General Education Board, 1919. 532 pp.

This is the last volume of the report of the survey of the Gary public schools and is based upon results of a number of educational tests given to the pupils in the Froebel, Emerson, Jefferson, and Beveridge schools between April 11 and June 9, 1916. The testing program was introduced by a series of nine simple psychological tests the scores from which are not included in the report. The purpose of giving these introductory tests was to familiarize the pupils with the procedure of testing. Each test was given throughout the four buildings on the same day except for occasional classes which were tested on the following day. Unusual care was taken to control the conditions of testing. Mr. Curtis and three trained assistants gave practically all of the tests. Automatic timers were used when feasible. The test papers were scored in duplicate. These precautions give the resulting scores a higher degree of reliability than they would have under ordinary conditions.

The tests and the grades to which they were given were as follows:

- I. Spelling.
 - Cleveland List Tests, grades two to twelve
 - Courtis Dictation Spelling Tests, grades two to twelve
 - Spelling was also tested by means of the Denver Composition Test, grades four to twelve
- II. Handwriting. The quality of the following types of material was measured by the Ayres' Handwriting Scale, "three slant edition."
 - Free Choice Handwriting Test, grades three to twelve
 - Denver Composition Test, grades four to twelve
- III. Arithmetic
 - Courtis Standard Research Tests, Series B, Forms 3 and 4, grades four to twelve
 - Cleveland Survey Arithmetic Tests, Sets C, G, H, L, and O, grades three to twelve
- IV. Reading
 - Kansas Silent Reading Test I, grades three to twelve
 - Kansas Silent Reading Test II, grades four to eight
 - Kansas Silent Reading Test III, grades eight to twelve
 - Gray's Oral Reading Test, grades three to eight
 - Reproduction Tests, grades two to twelve

Trabue Language Scales, B, C, D, and E, grades four to twelve

V. Language

Denver Composition Test, grades four to twelve

This volume stands out as distinctive in the series of volumes which make up the report of the survey of the Gary public schools. In the first place, it contains 532 pages besides the introduction and author's preface. The next largest, which is the summary volume, has only 265 pages. Second, it comes from the press several months after the others have been in the hands of readers and more than three years after the giving of the tests in Gary was completed. Third, the volume is a combination of the report of the survey and a treatise on the theory of educational measurements.

Besides a chapter given to each of the school subjects for which tests were given there are four other chapters: Chapter 1, "General Statement," in which is included a brief discussion of the status and limitations of educational measurements; Chapter 2, "Tests and Testing Conditions;" Chapter 8, "Factors Affecting Performance;" Chapter 9, "Conclusions." There is an effort to make the report as readable as possible. The presentation for each school subject is divided into three parts. First, there is a concise non-technical discussion of the significant aspects of the data secured. Second, there is a discussion of "all that is involved in the testing process." Third, in the appendices of the volume there is reproduced a number of the tabulations and a quantity of other data in which students of education will be interested. This organization materially facilitates the reading of the report for one who is not interested in critical discussions of the tests. One is disappointed, however, in not finding a concise summary statement of the findings for each subject. A partial exception to this statement should be made in the case of handwriting. A judicious use of paragraph headings would materially increase the clarity of the report.

The volume must be judged both as a report of the Gary survey and as a treatise on educational measurements. The conception of the survey is shown to have been comprehensive and the precautions exercised in its execution may well be taken as a model for future surveys. For example, several tests were used in each school subject. This is good procedure, because no single test can possibly have the same reliability that a battery of several tests has. Furthermore, no one test can cover the entire field of a school subject. Again, the scoring was done in duplicate by pupils and checked by examiners. This also contributes to the reliability of the published results. Among a number of other instances of the precautions taken, one may mention the fact that the judges who rated the compositions were carefully trained to their work before they began scoring.

In general the interpretation of the scores is made with unusual care and fairness. In English composition, for example, pupils wrote on the same subject and under the same conditions as obtained in the surveys of Denver and Grand Rapids. The compositions from these latter cities were, however, measured by the Willing Scale. In order to secure comparative data by which to interpret the Gary results it was obviously necessary to have the Gary papers rated on the same scale. The added precaution

was taken of having them (those of the eighth grade) rated by the same person—namely Mr. Willing—who rated the papers at Denver and at Grand Rapids. Many other instances might be given of the unusual care and fairness with which interpretations were made. Indeed, the author shows a disposition to set up and dispose of as many possible objections as he can. He seems to assume—with some truth we admit—that his conclusions will be assailed. His state of mind in this matter might be described as “nervous.”

The author states as a general conclusion, “that the product of the class room teaching of the fundamentals is, at Gary, poor in quality and inadequate in amount; it approximates in character the product of the poorer conventional schools, and reveals in no particular the slightest indication that it has been affected either favorably or unfavorably by the enriched curriculum or other special features of the Gary schools. . . . The entire investigation reveals many and consistent evidences of careless work, imperfectly developed habits and marked lack of achievement.” (p. 384) A careful study of the results of the tests reveals many instances which confirm the above statement. In many cases the Gary scores are below comparative data from conventional schools. In some cases, however, the Gary scores compare favorably with the “available standards” and in a few cases comparative data are not used which would result in a more favorable though not necessarily reversed conclusion in respect to Gary.

For example, in interpreting the handwriting scores for quality comparative scores from Iowa, Cleveland, St. Louis, Grand Rapids, and Kansas City and the standards proposed by Freeman and Starch are used. It is entirely possible that the average performances of the pupils in conventional schools are too high for satisfactory standards. No reference is made to the studies by Ashbaugh¹ and Koos² in which they indicate that possibly the standards for quality should be lower than Freeman and some other investigators have indicated. Ashbaugh shows that the New York Civil Service Commission accepts 40 and 50 for “passing” in writing; and Koos proposes an ultimate standard of 60, even finding “considerable justification” for 50. The eighth-grade median for Gary was 43.

The conclusion is reached that “handwriting instruction at Gary is producing very small effect upon the product.” (p. 46) This gives one the impression that the handwriting is entirely unsatisfactory. While it is true that quality at Gary is below even the lowest proposed standard, it is equally true that rate is notably above the rates used as comparative data. Rate of writing is important and credit should be given for this phase of the product of instruction in handwriting. Indeed, the author does so in other connections. It may be noted in passing that this is not the only instance in the report of inconsistency of statement.

¹ Ashbaugh, E. J. *Handwriting of Iowa school children*. (Bulletin of State University of Iowa, March, 1916.)

² Koos, L. V. “The determination of ultimate standards of quality in handwriting for the public schools,” *Elementary School Journal*, 18:423, February, 1918.

Another example may be cited from the chapter on English composition. In addition to judging compositions by means of the Hillegas Scale, a study was made of the vocabulary used by the eighth-grade children at Gary. Comparison is made with the vocabularies reported by Jones and the conclusion is reached that "a careful study of the vocabulary fails to show any clear effect of special training at Gary." In the discussion there is no recognition of the possible effect of the foreign population of Gary upon the vocabulary of the children. This factor may have been taken into consideration but there is no direct evidence of it. Neither does he consider the possible effect of the fact that each child at Gary wrote but one composition—and that on "an exciting experience"—while the children in Jones' study wrote on the average 105 compositions each and were judged to have exhausted their vocabularies.

A third example of conclusions at variance with the sweeping general statement quoted above may be taken from the chapter on reading. The Kansas Silent Reading Tests yielded median scores at Gary which approximate the standard for conventional schools from the sixth grade to the eighth grade inclusive and which are rather notably above standard for grades nine, ten, and eleven. It is only in grades three and four that the Gary scores are materially below standard. It is true that these facts are stated by the author in their appropriate places. The point, however, is that they are not reflected in his general conclusions.

In the summary volume the following paragraph appears. "The impressions gained through inspection and the results of the tests are not entirely consistent. For example, silent reading makes a distinctly better showing in the tests than one would have expected on the basis of class room observation. Spelling makes a poor showing on the list test and a very good showing on the composition test. It is not possible to reconcile these divergencies without adducing considerations as to which different opinions could fairly be entertained. Without, however, seeking to ignore the conflict of evidence, the authors still feel that the quality of class room instruction at Gary falls short of what is necessary." (p. 102)

After studying the survey report on "Measurement of Classroom Products" and finding the interpretation of results open to such criticisms as indicated above, one cannot help feeling that at times the author has allowed his interpretations to be colored by what the writers of the summary volume suggest to have been the prevailing opinion of the survey staff. Even if such instances are not numerous in the report, it is unfortunate that they exist at all, because the objective nature of standardized tests has probably been the most convincing argument in favor of their use. If the measurement is objective but the interpretation becomes subjective, this argument loses its force.

As a treatise on the theory of educational measurements, the volume is a notable contribution. It contains a discussion of the nature of the abilities and detailed descriptions and analyses of certain of the tests used. Mr. Courtis gives for the first time an account of the construction of the Courtis Standard Research Tests in Arithmetic, Series B. A more complete

account is given of the validity and reliability of most of the tests used than has heretofore been published.

In Chapter VIII the author considers the factors upon which performance depends. This constitutes an important analysis, and throws an unexpected light on the meaning which may be attached to the Gary scores in different tests. The basic factor upon which performance depends is heredity or "capacity." The maturity of the pupil is a second important factor. Maturity together with general training (a third factor) is shown to be more potent in the Gary schools than specialized training (a fourth factor) in the development of the following abilities: rate of copying figures, oral reading, writing of multiplication combinations, handwriting (all tests), addition in Series B, and rate of silent reading as revealed by the Kansas Silent Reading Test. On the other hand, specialized training is more potent in engendering accuracy of performance—e.g., in silent reading, long division, addition and subtraction of simple fractions. This distinction is fundamental in judging the efficiency of specialized instruction on the basis of scores yielded by standardized tests—and it is a distinction which Mr. Courtis has more ably presented than any other survey writer.

Students interested in the theory of educational measurements will find the report a source of much valuable information. The various "critical discussions" of the tests, however, which are announced to contain "all that is involved in the testing process," fall disappointingly short of the realization of this ambitious announcement. For example, in the critical discussion of spelling, the author points out that the ability to spell is "specific" (although it may contain a general factor), and that from a child's ability to spell one word no reliable inference can be made concerning his ability to spell any other word. However, a pupil's spelling of a number of words can be used as a basis for inferring his ability to spell other words belonging to the same difficulty group. The reliability of the inference is greater when random samples are used and increases with the number of words. Unfortunately, the discussion does not deal with the relation between the reliability of the inference and the number of words used, except to say, "it seems to the writer probable that the conclusions would not be changed were the number of test words to be largely increased."

Neither does the author investigate the relation between the reliability of a test and the difficulty of the words used, or the relation between the reliability of a test and the method of giving the test words. These omissions occur in spite of the author's statement that "the preceding discussions have had for their purpose the full statement in regard to the validity of the performance of an individual in a single test as a measure of his ability in spelling."

The discussions for the other subjects have similar limitations. At the present stage of the development of measuring instruments and the analysis of the measuring process, it is not to be expected that the "critical discussions" would be complete. In fact, they contain much more than one might expect, since they are somewhat incidental to a survey report. The propriety of attempting to include in such a report a discussion of "all that is

involved in the testing process" will doubtless be questioned by many readers. One finds many pages that appear to contribute very little to a report of the "measurement of class room products," although undoubtedly they will be commended by those who are interested in the theory of educational measurements and would very properly find place in a treatment of that topic. The inclusion of these topics undoubtedly delayed the preparation of the report, and in this way lessened the value of the volume as a report of the survey.

In the dual character of the report, there is evidence that as yet we have not found a satisfactory procedure for reporting the use of educational tests in a school survey. If the author gives no more than the results and his interpretation of them, he may be accused of misinterpretation and even misrepresentation by some of those who do not understand the process of measurement and the limitations of the particular tests used. On the other hand, if, as in this case, the author enters into a critical discussion of the tests and the theory of measurement, the report may lack unity and will certainly become voluminous. It may fail to be effective in making clear the actual conditions which were found to exist. Presumably Mr. Courtis was interested in the opportunities that the survey afforded for research. At any rate, it is certain that as a contribution to the literature of educational research his book is a more notable contribution than it is as a report of the measurement of classroom products at Gary.

W. S. M.

BUCKNER, CHESTER A. *Educational diagnosis of individual pupils* (Teachers College, Contributions to Education, No. 98). New York: Teachers' College, Columbia University, 1919. 93 pp.

FRETWELL, ELBERT K. *A study in educational prognosis* (Teachers College, Contributions to Education, No. 99). New York: Teachers' College, Columbia University, 1919. 55 pp.

These are companion volumes in the sense that the investigations which they report were carried out upon the basis of the same tests of the same boys.

Dr. Buckner investigates the methods by which scores in one test may be legitimately compared with scores in another, the amount and distribution of individual variability, the existence of erratic scores, the causes of extremely variant scores, and the relation between ability and variation. To express these scores in terms of the standard deviation or the quartile deviation permits comparison between them, and at the same time retains the refinement of the original scores.

As usual it was found that pupils varied widely in a given test, extreme scores being more numerous below the median than above it. It was also found that a given pupil differed markedly in his scores on different tests, the scores having first been made comparable. It appears that the more capable pupils—if we may draw practical conclusions from theoretical discussions—not only do better work but that they do it more consistently.

It is pointed out, moreover, that while some of the variability and some of the extreme scores are caused by distractions and other chance occurrences,

these factors are not as effective in producing high scores as in producing low scores. The practical meaning of this would seem to be that pupils who obtain extremely low scores are entitled to the benefit of further investigation.

It seems, therefore, to be clear that the determination of the status of a pupil must be carried out on the basis of a considerable amount of information concerning each pupil. This is the reason for teams or "batteries" of tests.

As to "Causes of Extreme Variability," the following are considered: (a) the nature of the tests; (b) their administration; (c) accidental occurrences; (d) statistical treatment of the results; and (e) the ability of the pupil in different traits.

Two important practical considerations as to the nature of tests are distinguished. First, the range of a test must be somewhat greater than the range of ability in the group tested. Second, the intervals between possible scores on the test must be small.

We shall not take up each of these five causes. Naturally it is the ability of the pupil which most affects extreme variability. One of the practical conclusions reached in this connection is that extreme scores are little, if any, more unreliable than those near the median. This appears to be particularly applicable to low scores despite the fact that accidental occurrences tend to depress rather than raise results. The evidence on this is the fact that ability and variability bear a direct relation to each other.

Dr. Fretwell takes up the investigation at the point where Dr. Buckner leaves it. He asks: "Can academic success be predicted? If so, how?" The conclusion is reached that academic success can be predicted. When all temporary illnesses, ranging from bad colds through contagious diseases to a month in the hospital, all fortunes or misfortunes in the home life, barring withdrawal from school, all the change in physical conditions and varying interests among boys of eleven to thirteen—when all these and a score of others that might be enumerated are considered—the use of these tests in one and one quarter hours' testing at the beginning of the year would have agreed with the classification of the teacher after teaching the pupils one year in 90 percent of all the cases.

As to the best basis of prediction, three others besides the tests are considered—school marks from the beginning of each pupil's school career, age, and the combined estimates of departmental teachers. The results of these comparisons lead the author to say: "In this study, academic success in the first year of junior high school was more successfully predicted by a group of standardized tests than by all previous school marks or age or teachers' estimates."

Some tests prove more valuable for prediction than others. The following is the order of importance as the author found it: reading (Thorndike's *Understanding of Sentences*), visual vocabulary, opposites, spelling, completion tests, arithmetic tests (Woody), easy directions, mixed relations, and composition.

B. R. B.

WILSON, GUY MITCHELL. *A survey of the social and business usage of arithmetic*. (Teachers College. Contributions to Education, No. 100). New York: Teachers College, Columbia University, 1919. 62pp.

The content of school subjects is largely determined by tradition. This simply means that the content as originally determined tends to be retained. Changes are made from time to time but in general they lag far behind social and school needs. Recently the attention of educators has been directed to curriculum making, and many investigations have been undertaken to determine what should constitute the curriculum in a number of school subjects.

Tradition has influenced to a very large degree the curriculum of arithmetic as it is now taught. Since the work of Rice in 1893 and McMurry's proposals enunciated in 1904, a number of investigators have attacked the problem of determining what the course of study in arithmetic should be. Professor Wilson has employed a procedure which is different from that used by preceding investigators.

He states that the "problem set for solution in this study is to determine the arithmetic actually used by adults in their social and business relations." Accordingly pupils in the sixth, seventh, and eighth grades in 23 cities, and in certain rural schools were requested to secure from their parents the arithmetical problems which the latter had solved during a period of two weeks. Reports were obtained from 4,068 persons and the total number of problems after unsatisfactory reports were rejected was 14,583. The author demonstrates that the group of persons reporting is probably sufficiently representative of the population of the entire country with the exception of communities which have specialized activities.

The 14,583 problems were analyzed as to the topics involved and the magnitude of the numbers occurring in the four fundamental operations. For both integers and fractions, common and decimal, the author found that six topics (namely, multiplication, addition, subtraction, division, fractions, and accounts) include slightly more than 90 percent of all the problems. The remaining 10 percent are divided among 37 additional topics.

With reference to the magnitude of the numbers, the author found that they rarely exceeded four places and that the majority of them did not extend beyond two places. He found only 50 different common fractions in all of the problems reported, and 40 of these occurred eight times or less.

Although the title and the explicit statement of the author's problem indicate that the study was made to determine what arithmetic is actually used by adults in their social and business relations, it appears from a reading of the monograph that the author's real purpose was to determine in a general way what should constitute the curriculum in arithmetic in the elementary school. In making this application of his results, he announces two main assumptions:

"(1) *Aim of Arithmetic in the Grades*. While not denying the cultural and disciplinary value of arithmetic—in common with any subject systematically studied and well taught—it is assumed that arithmetic in the grades is justified only on the basis of its utility in the common affairs of life. We learn the multiplication table not to sharpen the wits nor to comprehend a

beautiful system but to figure our bills, our taxes, or the interest on a note. Whatever arithmetic is given in the grades beyond the essentials required by social utility consumes time that could be used more profitably in other ways.

"(2) *Determining the Social Utility of Arithmetic.* Social utility can be determined only on the basis of the uses of arithmetic in adult activities."

Professor Wilson would eliminate decimal fractions except in United States money, proportion which he states is really absent, square root, reductions in denominate numbers, exchange, and the metric system in addition to eighteen other topics. He believes it to be safe and conservative to recommend the omission of these topics. He says, "If to the four fundamentals and fractions, one were to add accounts, simple denominate numbers, and percentage, little would be left for all the other processes. So little, in fact, that it seems unfair to give attention to them as drill processes in the elementary schools. Some of them should receive no attention.

It appears that the author defines the use of arithmetic as "figuring" and that three additional assumptions are implied: (1) that adults are now making the uses of arithmetic which they should make in their social and business activities; (2) that the needs for "figuring" will be approximately the same in the future as they are now for the adults included in this study; (3) that the best way to prepare a child for "figuring" is to use a curriculum based solely upon the problems which adults use.

The author does not consider the school needs which a pupil may have for arithmetic in the study of geography, history, science, manual training, and other school subjects. Until the curriculums for these subjects are changed radically, pupils will have need for certain aspects of arithmetic. In the absence of a scientific analysis of their arithmetical requirements, one should not assume that they are included in the adult social and business usages of arithmetic.

Doubtless many readers will be unable to agree with Professor Wilson in his assumption that no arithmetic should be taught in the elementary school beyond the "essentials required by social utility" and that "social utility can be determined only on the basis of the uses of arithmetic in adult activities." Doubtless most readers will be willing to accept this as one of the important bases; but absolutely to reject other bases—as the author does—is to take a radical position. This is especially true since he implicitly makes other assumptions, as has been pointed out, which deserve consideration in the making of a course of study.

To define the usable product of arithmetical instruction as "figuring" is to limit decidedly the significance of the subject. It is true that "figuring" is the most tangible and conspicuous outcome; but other outcomes, although they are more subtle, are nevertheless important and should be definitely planned for in the making of a curriculum. Furthermore, it is not safe to build a curriculum on the assumption that adults are now making all the use of arithmetic which they should make in their social and business activities. There seems to be considerable unanimity of opinion that in many respects adults are now lacking in efficiency with respect to quantitative matters. As an instance of this we may cite the agitation for family budgets and farm accounting.

Professor Wilson's study will no doubt have a far-reaching effect upon school practice. It should, because it is by far the most valuable work which has been done on minimum essentials in the field of arithmetic. As a survey of the social and business usage of arithmetic the monograph is a notable contribution and will probably rank, as the author suggests, with that of Ayres and Jones in the field of spelling. To the writer it appears, however, that he has lost sight of certain important considerations in applying the results of his survey to the making of a curriculum for arithmetic, as is indicated in the implied assumptions given above. Some of his conclusions with respect to the curriculum will probably stand when a more conservative and analytical interpretation is made of the data collected, but all of them should not be accepted until the effect of the implied assumptions and possibly of other factors is known.

W. S. M.

Bibliographies

STANDARDIZED TESTS FOR THE HIGH SCHOOL

(Continued from March number page 242)

C. Algebra (Continued)

Stromquist's Preliminary Algebra Tests. This series includes tests upon the following operations: (1) addition; (2) subtraction; (3) multiplication; (4) division; and (5) factoring.

Publisher: University of Wyoming, Laramie, Wyoming.

Thorndike's Algebra Test. This is a series of eight exercises arranged in order of increasing difficulty as determined by the opinion of competent judges.

Reference: Thorndike, E. L. "An experiment in grading problems in algebra," *Mathematics Teacher*, 6:123-34, March, 1914.

D. Geometry

Irwin's Tests in Mental Manipulations of Space Relations. The author feels that this ability is not confined to mathematics alone, but it is constantly in demand in other lines as sewing, general science, drawing, English, etc. He has, therefore, worked out four tests (Tests A-D) which should "yield a rough measure of the ability of each pupil in a class to image space relations."

Address: H. N. Irwin, Fairmount Junior High School, Cleveland, Ohio.

Reference: Irwin, H. N. "Preliminary attempt to devise a test of the ability of high school pupils in the mental manipulations of space relations," *School Review*, 26:600-6, 654-70, 759-72, October-December, 1918.

The derivation of the tests is described and the tests are reproduced.

Minnick's Geometry Tests. This series of tests is based on the assumption that demonstration of a geometrical theorem involves the following

abilities: (1) the ability to draw the figure; (2) the ability to state the hypothesis and conclusion; (3) the ability to recall facts concerning the figure; (4) the ability to select and organize facts so as to produce the proof.

Publisher: J. H. Minnick, University of Pennsylvania, Philadelphia, Pennsylvania. Also obtainable from Bureau of Educational Research, University of Illinois, Urbana, Illinois.

Reference: Minnick, J. H. "A scale for measuring pupils' ability to demonstrate geometrical theorems," *School Review*, 27:101-9, February, 1919.

Study based on tests given in thirty high schools throughout the country and ranging in size from a few hundred pupils to several thousand. These pupils had completed either the first two books of plane geometry or all of plane geometry. Illustrated by graphs and tables.

Minnick, J. H. "Certain abilities fundamental to the study of geometry," *Journal of Educational Psychology*, 9:83-90, February, 1918.

The author discusses the four abilities which are fundamental to the study of geometry; how are these abilities related to teaching; to what extent are these four abilities developed; development of a series of tests which may be used for the purpose of diagnosis.

Minnick, J. H. *An investigation of certain abilities fundamental to the study of geometry*. Philadelphia: University of Pennsylvania, 1918.

In this monograph Dr. Minnick gives an analysis of geometrical ability and a complete account of the derivation of his tests.

Starch's Geometry Test.

Address: Daniel Starch, University of Wisconsin, Madison, Wisconsin.

Stockard and Bell Geometry Test. A test of seventy questions which involve drawing figures, naming figures, indicating order of development in demonstration, completing statements, stating the converse, definitions, regular polygons, parts of a demonstration, etc.

Stockard, L. V. and Bell, J. Carleton. "A preliminary study of the measurement of abilities in geometry," *Journal of Educational Psychology*, 7:567-80, December, 1916.

Derivation, preliminary form reproduced. Analysis of results.

IX. MENTAL TESTS

Army Tests. The mental tests which have been used extensively in the army were prepared by the Psychology Committee of the National Research Council. Three systems of tests are now in use:

1. *Alpha.* This is a group test designed for men who can read English. It requires only fifty minutes to give. The exercises are designed so they may be answered without writing, merely by underlining, crossing out, or checking. The papers may be scored by means of stencils, so that nothing is left to the personal judgment of those who do the scoring.
2. *Beta.* This is a group test for foreigners and illiterates. Success in Beta does not depend upon knowledge of the English language, as the instructions are given entirely by pantomime and demonstration.
3. *Individual Tests.* Three forms of individual tests are used: Yerkes-Bridges Point Scale, the Stanford-Binet Scale, and the Performance Scale.

Address: Division of Psychology, Medical Department, War Department, Washington, D. C.

References: *Army mental tests, methods, typical results, and practical application.* Washington, D. C., November, 1918.

This bulletin describes the three types of tests used in the army. It also gives an explanation of letter ratings and directions for the use of intelligence ratings.

Madsen, I. N. and Sylvester, R. H. "High school students' intelligence ratings according to the alpha army test," *School and Society*, 10:407-10, October 4, 1919.

Results of the Alpha test given to the high-school students of Rockford, Illinois, Madison, Wisconsin, and Sioux City, Iowa.

Van Wagenen, M. J. "Our schools as measured by the army tests," *Educational Administration and Supervision*, 5:163-76, April, 1919.

Mr. Van Wagenen states: "The use of the Army Tests in certain elementary schools, high schools, and universities for the purpose of gathering information for the use of psychologists in the army has, at the same time, revealed very clearly certain facts about the schools, the students tested, and possible values of mental tests in educational practice." The author describes Army Test E and compares it with Alpha, form 6. Table II gives the weighted scores of elementary school pupils, high school freshmen, and college men and women.

Materials from the Army Tests under the title of "National Intelligence Tests" have been standardized by a committee of the National Research Council. They are to be published this spring, 1920, by the World Book Company, Yonkers, New York.

Holley's Sentence Vocabulary Scale. This scale has been devised with the Terman Stanford-Binet Vocabulary as a basis. The words of this well-known vocabulary are put in sentences, the last word of which is found among four words to the right. The pupils indicate the correct responses by underlining the word which completes the sentence. This scale has been found to be a fair measure of intelligence. Series 3B is for grades VII to XII.

Publisher: Bureau of Educational Research, University of Illinois, Urbana, Illinois.

Reference: *First annual report of Bureau of Educational Research*, 1918-1919.

Otis' Group Intelligence Scale. This scale is composed of ten tests which measure various phases of mental ability. Two forms of the scale are available. They may be used in grades VI to XII.

Publisher: World Book Company, Yonkers, New York.

References: Otis, A. S. "An absolute point scale for the group measurement of intelligence," *Journal of Educational Psychology*, 9:239-61, 333-48, May-June, 1918.

Report of the Department of Research, being part of a report of the superintendent of schools of Oakland, California, 1917, p. 210.

Theisen and Fleming's Classification Test. A group test based on the Army Test (Alpha), has been successfully used for high-school purposes. Takes fifty minutes to give and may be scored at the rate of 15 papers per hour.

Publisher: Bureau of Publications, Teachers College, Columbia University, New York City.

Reference: An article by Dr. Theisen entitled, "Group Intelligence Tests," will appear in an early number of the *Journal of Educational Research*.

Thurstone's Psychological Examination for College Freshmen and High School Seniors.

Address: L. L. Thurstone, Carnegie Institute of Technology, Pittsburgh, Pennsylvania.

X. MUSIC

Baldwin's Public School Music Test. A preliminary test of the efficiency of public school music. It is aimed to measure the accomplishment of pupils. In this respect it is similar to the Curtis Standard Research Tests in Arithmetic.

Reference: Baldwin, Ralph. "Efficiency in school music teaching and practical test of the same," *Music supervisors' national conference proceedings*, 1914, p. 43-50.

On page 46 the test is reproduced.

Seashore's Musical Talent Chart. This chart is based upon the analysis of musical ability and offers a graphic means of representing the pupil's musical ability. The pupil's musical abilities are measured by his response to tones which are mechanically produced. The chart is simply a convenient device for recording the responses of the pupil. The records for testing a child are made by the Columbia Graphophone Company, Woolworth Building, New York. The *Manual for Measures of Musical Talent*, furnished free by the Educational Department of the Columbia Graphophone Company, contains directions, norms, and interpretations.

References: Seashore, C. E. "The measurement of musical talent," *Musical Quarterly*, 1:129-48, January, 1915.

A discussion of the elements which make up musical talent and the derivation of the Musical Talent Chart, the chart itself, and its meaning and use.

Seashore, C. E. "The rôle of a consulting supervisor in music," *Eighteenth Yearbook of the National Society for the Study of Education*, 1919, part II, p. 111-23.

Account of the use of the test in the elementary schools and with adults. Mr. Seashore states that the best time to use the chart is in the fifth grade.

Seashore, C. E. *The psychology of musical talent*. Boston: Silver Burdett and Company, 1919.

The system of evaluation of musical talents is reviewed and a chapter is devoted to each of the five double disk records.

Seashore, C. E. *Vocational guidance in music*. (University of Iowa Monographs, first series, no. 2, September, 1916.)

Copies may be obtained on request from the librarian of the University of Iowa, Iowa City, Iowa. A discussion of the value of the Musical Talent Chart.

"Testing for musical talent," *Musician*, 24:10, July, 1919.

A discussion of Seashore's tests.

XI. PHYSICAL TRAINING

Athletic Badge Tests for Boys. These tests are not a complete measure of physical efficiency, but they serve as a fair index of heart, lung, and muscular development. They comprise a sixty-yard dash, standing broad jump, and a pull up. Standards have been worked out. These tests have been used in the high schools of Gary, New Orleans, Seattle, Buffalo, and New York.

Publisher: Playground and Recreation Association of America, 1 Madison Avenue, New York City.

Reference: Hanmer, L. F. *Gary public schools: physical training and play.* New York: General Education Board, 1919.

Athletic Badge Tests for girls. This test is similar to the one for boys, but being a newer series has no standard.

Publisher: Playground and Recreation Association of America, 1 Madison Avenue, New York City.

Reference: Hanmer, L. F. *Gary public schools: physical training and play.* New York: General Education Board, 1919.

Mr. Hanmer states: "The results (for the girls) are therefore no more satisfactory than the results of the boys' test. There is also the same unevenness and irregularity in development."

Baldwin's Physical Development Scale. Separate scales are given for boys and girls. Each scale gives norms for height, weight, and breathing capacity for each year from the kindergarten to the last year of the high school. This scale is included in Rapeer's Scale for Measuring Physical Education.

Address: Bird T. Baldwin, University of Iowa, Iowa City, Iowa.

Reference: Baldwin, B. T. "A measuring scale for physical growth and physiological age," *Fifteenth Yearbook of the National Society for the Study of Education*, 1916, part I, pp. 20-21.

Rapeer's Scale for Measuring Physical Education. This is a score card for judging five aspects of the results of physical education, i.e., health, physiological efficiency, physical development, physical ability and mental qualities.

Address: L. W. Rapeer, 1719 H. Street N. W., Washington, D. C.

XII. READING

Kansas Silent Reading Tests. Test III of this series is designed for the high school. It consists of a series of exercises in which the pupil is asked to read a short paragraph and answer a single question on it. The administration of the test is very simple and it has been very widely used.

Publisher: Bureau of Educational Measurements and Standards, Kansas State Normal School, Emporia, Kansas.

References: Kelly, F. J. *The Kansas Silent Reading Test* (Studies by the Bureau of Educational Measurements and Standards, No. 3), Emporia, Kansas: Kansas State Normal School, 1915.

This contains the complete account of the derivation of the test.

Kelly, F. J. "The Kansas silent reading test," *Journal of Educational Psychology*, 7:63-80, February, 1916.

This is a partial account of the derivation of the test.

Monroe, Walter S. "A report on the use of the Kansas silent reading tests with over one hundred thousand children," *Journal of Educational Psychology*, 9:600-8, December, 1917.

The author concludes that the Kansas tests place more emphasis upon comprehension than upon rate of reading.

Smith, B. M. "Correlation of ability in reading with the general grades in high school," *School Review*, 27:493-511, September, 1919.

In trying to determine the correlation between abilities in reading and school abilities in the Oak Park, Illinois, high school Miss Smith used the Kansas Silent Reading and three other tests which she devised. Correlations are given for Kansas Silent Reading as well as the other tests.

An educational survey of Janesville, Wisconsin. Madison, Wisconsin: State Department of Public Instruction, 1918, p. 261.

An account of the use of the Kansas Silent Reading tests in high school.

Monroe's Standardized Silent Reading Tests. This series of silent reading tests is essentially a revision of the Kansas Silent Reading Tests described above.

Publisher: Bureau of Educational Research, University of Illinois, Urbana, Illinois.

Reference: Monroe, Walter S. "Monroe's standardized reading tests," *Journal of Educational Psychology*, 9:303-12, June, 1918.

This is an account of the derivation of the revision. The Kansas Silent Reading Tests have been criticized because many of the tests resembled arithmetical puzzles rather than ordinary reading material. In the revision the paragraphs have been selected from school readers.

School report of the town of Southington, Connecticut, 1919, p. 17.

An account of the use of these tests in the high school.

Neher's High School Vocabulary Scale. A test consisting of one hundred words chosen from the top of every third column of Laird and Lee's *Vest Pocket Dictionary*.

Address: H. L. Neher, Northwestern University, Evanston, Illinois.

Reference: Neher, H. L. "Measuring the vocabulary of high school pupils," *School and Society*, 8:355-59, September 21, 1918.

Reproduction of the test, scores obtained through experimentation and conclusions reached.

St. Paul Survey Silent Reading Scale for High School. This scale is similar to Minnesota Scale Beta, the derivation of which is based on Thorndike's Scale Alpha 2.

References: *Report of a survey of the school system, St. Paul, Minnesota*, 1917, p. 375-401.

The tests and questions are reproduced. Methods of scoring and a comparison of scores is also given.

Haggerty, M. E. and Thomas, M. J. "Preliminary study of the reading attainments of college freshmen," *School and Society*, 6:230-38, August 25, 1917.

The test was given to the college freshmen and was also given to two hundred and fifty members of the senior class in the St. Paul high school. The test which is reproduced in this article includes a number of the exercises from the St. Paul Survey Test.

Smith's Silent Reading Paragraphs for High Schools. Miss Smith has devised three informal silent reading tests. The tests are as follows:

Test I. Paragraph Test. Ability to grasp central idea. Four paragraphs were chosen from "A Father to His Freshman Son in College," (*Atlantic Monthly*).

Test II. Information Test. Ability to recall items of information. Selections were taken from the *Independent*.

Test III. Narrative Test. Ability to follow plot. Selection was a story from an inside page of *Chicago Daily Tribune*.

Address: Bertha M. Smith, Oak Park High School, Oak Park, Illinois.

Reference: Smith, B. M. "Correlation of ability in reading with general grades in high school," *School Review*, 27:493-511, September, 1919.

The tests were given to determine the correlation between abilities in reading and school abilities. The four reading tests used are described and the author's three tests are reproduced.

Starch's English Vocabulary Test. This test is designed to measure the percent of words of the entire English vocabulary as well as the absolute number of words that a person understands. The test consists of several sets of one hundred words which were selected at uniform intervals from the entire English vocabulary—Webster's *New International Dictionary*.

Address: Daniel Starch, University of Wisconsin, Madison, Wisconsin.

Reference: Starch, Daniel. *Educational measurements*. New York: Macmillan, 1916, p. 38 ff.

The test is reproduced. Tentative scores are shown for different grades in the high school as well as for elementary school and college.

Thorndike's Scale Alpha 2 for Measuring the Understanding of Sentences.

This scale consists of a series of paragraphs and questions about the paragraphs arranged in order of difficulty of comprehension. This scale measures only the degree of comprehension and does this in terms of difficulty of the exercise which the pupil is able to do 80 percent correctly. Part II may be used in grades VI to XII.

Publisher: Bureau of Publications, Teachers' College, Columbia University, New York City.

References: Thorndike, E. L. "Improved scale for measuring ability in reading," *Teachers College Record*, 16:445-67, 17:40-67, November, 1915, January, 1916.

The test, key and score sheet are reproduced in the November issue along with the discussion of the use of the scale method of scoring, and the interpretation of scores. The derivation of the test is in the November number.

Kelley, T. L. "Thorndike's reading scale alpha 2 adapted to individual testing," *Teachers College Record*, 18:253-60, May, 1917.

Directions for using the scale with individual rather than with the class as the author intended.

XIII. SCIENCE

A. General

Caldwell's Science Tests. The tests may be divided into three groups.

Group I consisting of two pairs of tests deals with observation and discrimination. Group II relates primarily to the pupil's ability to recog-

nize natural phenomena, to recall experiences in science study and to make constructive statements regarding these experiences. Group III consists of two tests of ability to give reasons for, to interpret, or to explain changing phenomena—the changes occurring in the presence of pupils in the form of an experiment or demonstration. These tests were used in the senior classes of the Gary, Indiana high school.

Reference: Caldwell, O. W. *The Gary public schools: science teaching.* New York: General Education Board, 1919.

Chapter X is devoted especially to the reproduction of the tests which have been prepared by Mr. Caldwell.

Downing's Information Tests in Science. A vocabulary test of one hundred scientific terms covering the entire field of high-school science. The pupils are asked to put the letter "E" before a word he can explain, and the letter "F" before a word which he has heard about.

Address: Elliot R. Downing, University of Chicago, Chicago, Illinois.

Reference: Downing, E. R. "A range of information tests in science," *School Science and Mathematics*, 19:228-33, March, 1919.

The author reproduces the list of one hundred terms and states results with about fifteen hundred students in seventh and eighth grades, high school, and university.

Grier's Range of Information Test in Biology.

Reference: Grier, N. M. "Range of information test in biology," *Journal of Educational Psychology*, 9:210-16, 388-93, April-September, 1918.

The design of the tests is that described by Whipple in his Range of Information Test. The first number contains a description of a test in physiology and the second one in zoology.

Herring's Tests in Scientific Thinking. Series of three tests on each of eleven abilities involved in scientific thinking.

Reference: Herring, J. P. "Measurements of some abilities in scientific thinking," *Journal of Educational Psychology*, 9:535-58, December, 1919.

The tests are reproduced and results of their application to one pupil is given.

Ruch's Range of Information Test in General Science. This test consists of fifty scientific terms assembled upon the basis of frequency in general science texts. As three series of words are given, a different series may be given during the year.

Address: G. M. Ruch, School of Education, University of Oregon, Eugene, Oregon.

Reference: Ruch, G. M. "A range of information test in general science," *General Science Quarterly*, 4:257-62, November, 1919.

Describes the derivation of the test which is reproduced. The method for scoring is discussed.

B. Chemistry

Bell's First-Year Chemistry Test. A test of 25 questions which the author considered basic for elementary chemistry.

References: Bell, J. C. "Study of the attainments of high school pupils in first-year chemistry," *School Science and Mathematics*, 18:425-32, May, 1918.

Brief account of its derivation. Results from 401 pupils analyzed.

Bell, J. Carleton. "A test in first year chemistry," *Journal of Educational Psychology*, 9:199-209, April, 1918.

This is the complete report of which the above is a preliminary statement.

Jones' Chemistry Tests. (Union Science Series, *See Physics*.)

Address: F. T. Jones, 10109 Wilbur Avenue, Cleveland, Ohio.

Rivett's Time Limit Test in Chemistry. The article presents two types of simple tests which are arranged for beginners. They are intended as suggestive types of tests which teachers of chemistry may easily devise. The pupils may score these tests themselves.

Address: B. J. Rivett, Northwestern High School, Detroit, Michigan.

Reference: Rivett, B. J. "Testing results in chemistry," *School Science and Mathematics*, 19:742-45, November, 1919.

Two types of tests are reproduced. Advantages over the old form examination are stated in conclusion.

Webb's Preliminary Test in Chemistry. Test requires pupils to classify fifty well-known substances as elements, compounds, or mixtures.

Address: Hanor Webb, George Peabody College for Teachers, Nashville, Tennessee.

Reference: Webb, Hanor. "A preliminary test in chemistry," *Journal of Educational Psychology*, 10:36-43, January, 1919.

This is an account of the use of the author's test with a number of his classes.

C. Geography

Boston Geography Test. A forty-minute test on United States and the countries of Europe. Although the tests were given to elementary, high, and normal school pupils, particular interest centers in the results achieved by the eighth grade.

Address: Leonard O. Packard, Boston Normal School, Boston, Massachusetts.

Reference: *Geography. A report on a preliminary attempt to measure some educational results.* (Bulletin No. 5 of the Department of Educational Investigation and Measurement, Boston Public Schools, 1916.)

Discussion of the tests with reproduction, but method of giving and of rating is not shown. Results for high-school students are given.

D. Physics

Chapman's Physics Test in Electricity and Magnetism, Sound and Light.

Address: J. Crosby Chapman, Western Reserve University, Cleveland, Ohio.

Reference: Chapman, J. C. "The measurement of physics information," *School Review*, 27:748-49, December, 1919.

The test with answers is reproduced. Results for 158 high-school pupils are tabulated.

Jones' Union Science Tests. These tests form an elaborate series covering many of the topics of the subject. There is a separate test for each topic.

Address and send results to: Franklin T. Jones, 10109 Wilbur Avenue, Cleveland Ohio.

Reference: Jones, F. T. "Practice exercises in physics," *School Review*, 26:341-48, May, 1918.

Describes the Union Science Tests—a series in physics. Test H₂—Test H₃, and Test H₄ are reproduced, with results analyzed.

Randall, Chapman and Sutton's Test in Mechanics. A simple test of a set of progressive problems in mechanics.

Address authors: Western Reserve University, Cleveland, Ohio.

Reference: Randall, D. P., Chapman, J. C. and Sutton, C. W. "The place of the numerical problem in high school physics," *School Review*, 26:39-43, January, 1918.

The test is reproduced. Results from experimental testing with four high schools are given.

Starch's Physics Test. This test is composed of 75 sentences which are arranged in the form of a completion test. The degree to which the pupils have mastered the work in physics covered by these sentences is shown by their ability to fill in the missing words, phrases, and numbers. The test is broken up into topics and standard scores for each topic are printed on the test sheet. There are 28 sentences on mechanics, eight on heat, nine on light, and 21 on magnetism and electricity.

Publisher: Daniel Starch, University of Wisconsin, Madison, Wisconsin. Also obtainable from Bureau of Educational Research, University of Illinois, Urbana, Illinois.

Reference: Starch, Daniel. *Educational measurements*. New York: Macmillan, 1916, chapter XIV.

XIV. VOCATIONAL SUBJECTS AND MANUAL TRAINING

Leavitt's Preliminary Test for Manual Arts. As the title suggests this is only a preliminary test, not designed by the author for general use.

Address: F. M. Leavitt, Associate Superintendent in Charge of Vocational Education, Pittsburgh, Pennsylvania.

Reference: Leavitt, F. M. "Standardized measurements in the field of industrial arts," *Industrial Arts Magazine*, 8:132-8, April, 1919.

Brief discussion of methods of producing and using standardized tests is given. The preliminary scale for manual arts is reproduced.

Wardner's Test for Knowledge of Tools. This is a modification of one of the Army Trade Tests. Pupils are asked to identify 81 tools. Test is reproduced.

Address: C. A. Wardner, Director of Cooperative Trade Schools, Springfield, Vermont.

Reference: Wardner, C. A. "Applying the army trade tests to vocational schools," *Industrial Arts Magazine*, 8:402-03, October, 1919.

A reproduction and brief discussion of this test which is being used in the Cooperative Trade School, Springfield, Vermont.

(The End)

News Items and Communications

National Intelligence Tests

In the spring of 1919 the General Education Board appropriated \$25,000 to the National Research Council to be used for the preparation of methods of measuring the intelligence of children in the elementary schools. A committee consisting of Dr. M. E. Haggerty, Dr. L. M. Terman, Dr. E. L. Thorndike, Dr. G. M. Whipple, and Dr. R. M. Yerkes, Chairman, was organized to undertake this task.

This committee has prepared two scales, the tests of which are listed below.

Tests of Scale A

Arithmetical Reasoning
Sentence Completion
Logical Selection
Synonym-Antonym
Symbol-Digit

Tests of Scale B

Computation
Information
Vocabulary
Analogies
Comparison

Either of these scales may be used alone, but the committee expects to recommend re-examination by the second scale after an interval of at least one day, in order that the reliability of the intelligence measurement may be increased. It is deemed preferable by the committee to use two short scales in the way suggested, rather than to use a single longer scale, the application of which would require approximately an hour of continuous work. Examination by either of the proposed scales, A or B, will require approximately thirty minutes.

Materials have been prepared by the committee for ten alternative forms of each of the ten tests which have been selected for use, but it is planned to complete for immediate publication only five forms of each test and of each scale. The remaining materials will be held subject to later development as needed.

It has been arranged that the World Book Company of Yonkers, New York, shall publish these new tests under the title "National Intelligence Tests." It is expected that the materials, which will consist of record blanks, examiner's guide, and scoring keys, will be ready for distribution in the spring of 1920. Royalties from the publications will be used for further research on methods of measuring intelligence and the practical application of such methods.

ROBERT M. YERKES
Chairman

The conclusions of E. J. Ashbaugh concerning the necessity of teaching derived forms in spelling¹ have been confirmed by the results obtained in tests given by the Department of Educational Investigation and Measurement, Boston, Massachusetts. While his data are based on tests given in the lower elementary grades (II to V), the data of this department are based on results obtained in grades VI and VII. There is also this difference in the two investigations. The words tested by Ashbaugh may or may not have been studied by the pupils since the tests were not organized on that basis, while in the Boston tests the root words given were specifically chosen from the Boston minimum spelling list which the pupils had studied.

Undoubtedly the adding of a suffix to a word makes that word harder to spell and we have no right to presuppose that pupils can spell derived words the roots of which they are able to spell. This is shown con-

COMPARISON OF ROOTS AND DERIVATIVES
(All root words occur in sixth-minimum list)

| Root Words | Percent Correct | Derivatives | Percent Correct |
|------------|-----------------|-------------|-----------------|
| abbreviate | 78 | abbreviates | 75 |
| absence | 85 | absences | 75 |
| accidents | 90 | accidents | 77 |
| arrange | 92 | arranging | 86 |
| attack | 94 | attacked | 87 |
| | | attacking | 94 |
| boundary | 87 | boundaries | 67 |
| century | 93 | centuries | 77 |
| commence | 85 | commenced | 81 |
| | | commencing | 74 |
| conquer | 89 | conquered | 80 |
| deceive | 80 | deceiving | 70 |
| decide | 87 | decided | 85 |
| | | deciding | 76 |
| describe | 82 | describing | 76 |
| industry | 96 | industries | 79 |
| judge | 98 | judges | 93 |
| material | 84 | materials | 75 |
| mosquito | 94 | mosquitoes | 59 |
| recognize | 87 | recognized | 85 |
| return | 98 | returned | 93 |
| | | returning | 95 |
| salary | 96 | salaries | 77 |
| telegraph | 98 | telegraphed | 90 |
| telephone | 98 | telephoned | 95 |
| ventilate | 87 | ventilated | 85 |
| wharf | 99 | wharves | 69 |

¹ Reported at February 1919 meeting of the National Association of Directors of Educational Research and in *Journal of Educational Psychology*, March, 1919, p. 143.

clusively in the results of spelling tests given in Boston in 1916,² 1917, and 1919. In each of the years 1916 and 1917 a seventh-grade division of approximately 1,000 pupils was given a list of root words to spell; another seventh-grade division was given derivatives of these same words; i.e., the first division spelled such words as "health," "accurate," and "promise," while the second division spelled "healthy," "accurately," and "promised." With few exceptions there was a falling off, and in some cases a very large falling off, in the percent of accuracy with which the derived forms were spelled, even in cases where the difference in spelling was only the addition of "s" or "ing" with no change of the existing root.

As a result of these findings the teachers were urged to teach in their spelling lessons the more common derivatives in connection with the root word.

In the spelling test of May 1919, 35 sixth-grade classes spelled the following list of 23 root words (all included in the sixth-grade minimum list) and 27 derivatives of those words with the percent of accuracy indicated in the table on page 322.

In every case but one ("attacking") the derivative proved itself a harder word to spell even though the teachers had been urged to teach the common derivatives. Since it is as important that the pupils know how to spell such common words as "judges," "mosquitoes," etc., as well as the root words, *every common derived form must be considered as presenting a new spelling problem and must be taught as such.*

HARRIET M. BARTHELMSS

Conference on Educational Measurements The Seventh Annual Conference on Educational Measurements will be held at Indiana University, Friday and Saturday, April 23 and 24. Indiana University was the first institution to hold an annual conference on educational measurements. These meetings are well attended by the school men of the state; and the university has brought to this meeting year after year speakers of national reputation. The committee in charge of the conference this year has issued the following preliminary announcement:

"The principal speakers for the Seventh Annual Conference on Educational Measurements will be Dr. Edward L. Thorndike of Teachers College, Columbia University, and Professor C. E. Seashore of the University of Iowa. Dr. Thorndike's discussions will be along the general lines of intelligence examinations for college entrance, standardization of textbooks, and the dangers of indiscriminate use and misuse of educational tests. Professor Seashore, who has done some remarkable work in the particular field he will present, will bring to the Conference the results of his research in the measuring of special abilities, particularly the measuring of ability in music."

² See Bulletin XI of the Department of Educational Investigation and Measurement, Boston, Massachusetts.

A number of different types of tests have been devised for measuring the silent reading vocabulary of pupils. The best known of these probably is the Thorndike's Visual Vocabulary Scales A2 and B. In these the pupil is asked to indicate the class to which the word belongs as flower, animal, or boy's name. In other vocabulary tests, the pupil is asked to check the words he knows the meaning of and to write out the meaning of those concerning which he is uncertain. Both of these plans have obvious limitations. The Southington-Plymouth English Vocabulary Scale devised by Ernest C. Witham makes use of a different device. Fifty English words are given. On the opposite page their brief definitions as given by Webster's *International Dictionary* are printed. The pupil is to record before each definition the number of the word which it fits.

Most of the words were selected from Ayres Spelling Scale. A few additional words were taken from Thorndike's Visual Vocabulary Scale A2, series α . The test has been tentatively standardized for the eighth grade and the high school. Copies of the test can be secured by addressing Superintendent Ernest C. Witham, Southington, Connecticut.

Mr. George N. Cade, Supervisor of Elementary Grades, Springfield, Illinois, gave Monroe's Standardized Silent Reading Test II to the VIII^A pupils who were promoted to high school at the end of the first semester. The principal of the high school has classified these pupils in the English department on the basis of their ability to read silently as shown by this test. Since ability to read silently is fundamental to practically all of the work of the high school, classification of pupils on the basis of this ability may be expected to effect a helpful organization.

The Myers Mental Measure by Caroline E. Myers and Captain Garry C. Myers, claimed by the authors to be the first group intelligence test applicable to all ages, is recommended for use in every grade since the directions are given orally and since the exercises consist wholly of pictures. It can be given to illiterates as well as to those who can read. The time for giving it is not more than twenty minutes. The publisher is The Sentinel, Carlisle, Pennsylvania.

Professor M. E. Haggerty of the University of Minnesota devised for use in the Virginia School Survey two tests of general intelligence, Delta 1 for grades one to three and Delta 2 for grades three to nine; and a test for measuring achievement in reading, Sigma 1 for grades one to three.

These tests have recently been published and made available for general use by the World Book Company of Yonkers, New York. The tests are accompanied by a manual of directions and keys for scoring. Their use in the Virginia Survey and in a number of cities elsewhere in the country insures their standardization.

National Association of Directors of Educational Research

(E. J. ASHBAUGH, *Secretary and Editor*)

Even though it is several weeks since the meeting of the Department of Superintendence at Cleveland, this is the first opportunity to tell you, who were not fortunate enough to attend, something of the meetings of our association at that time.

The first annual closed meeting on Monday was a decided success. Although conflict of time with the meetings of the College Teachers of Education prevented a few members from attending, the number present was quite satisfactory and the program even more so. Though announced as a closed meeting, nothing short of a corporal's guard would have kept some interested non-members out. It was more than curiosity on their part, since once inside they remained throughout the meeting.

The forenoon was given over to the report of the Committee on Standards (Mr. Curtis, Chairman) and a paper on "Diagnosis" by Mr. Kallom of Boston, the paper being presented by Miss Barthelmess in his absence. The report of the Committee on Standards attempted to accomplish its task by: (1) pointing out the problems to be solved, the ends to be achieved; (2) indicating so far as is at present possible the criteria to be used in judging of values; (3) providing an incentive to induce workers to apply these criteria; and (4) leaving it to a process of natural selection to determine which method shall survive. The association plans to make available to all its members as well as to the public the terms, methods, and practices approved. The association passed a motion for the reorganization of the committee with provision for subcommittees in a number of special fields.

The paper on "Diagnosis" set forth the method the Boston department is using in acquainting the assistant superintendents and masters with the position of classes relative to city standards. Although, as a result of these reports, considerable effectiveness is being registered in the teaching; it was admitted by the speaker that as yet the department has been unable on account of lack of resources to accompany its reports to the teachers with suggestions as to remedial measures.

At 12:30 luncheon was served in "our" private dining room. Thirty of the members and their guests participated.

The afternoon program included two papers on "The Evaluation of Tests"; one by Dr. Monroe and the other by Dr. T. L. Kelley. Both speakers led the members into a very deep discussion of the need for the statistical methods involved in the determination of the validity and reliability of tests. It

involved a consideration of technic impossible for the majority of us to practice but represented a phase of the testing work in which far too little has been done.

President Buckingham gave the members a brief history of the evolution of our official organ, the JOURNAL OF EDUCATIONAL RESEARCH. We were delighted to learn that it is exceeding the hopes of its founders, that it is already on a self-supporting basis, and that the first issue has to be reprinted to supply the demand for it.

At six o'clock we again broke bread together in the comfortable adjoining room which the hotel management had assigned to us. By this time we felt as if we owned it.

On Thursday a program consisting of eleven numbers was presented. Judging from the attendance and the comment of those who heard our offerings, the program was a success. Most of the papers will be published in the JOURNAL OF EDUCATIONAL RESEARCH.

President Brown of Oshkosh led off with his paper on "Formulation of Method in Reading in the Light of Recent Investigations." As the chairman said in his introduction, Mr. Brown is a conspicuous example of a school man who has become an executive without ceasing to participate in research work. The morning program was organized about the topic of reading, and Mr. Brown's paper afforded a good introduction to the work of the morning.

Dr. Theisen handled the subject of "Provisions for Individual Differences in the Teaching of Reading" in an effective manner. He has gathered material from more than one hundred teachers on the provisions which they are making and the provisions which they feel ought to be made for adjusting the teaching of reading to differences in ability among children. The report was interesting because it showed just what the teachers are thinking about on the important matter of fitting instruction to the needs of pupils.

Mr. Kallom was to have presented a paper on "Reproduction as a Measure of Ability." He was, however, at the last minute, obliged to remain at home on account of illness in his family. We were sorry indeed not to greet Kallom but his paper was so capably presented by Miss Barthelmess, his assistant at the Boston Department of Research, that we almost forgot our disappointment. The point of the paper was that children reproduce what they have read in terms of their experience. Some of the responses of children were interesting not only from a scientific point of view but also were worthy of a place among the "bright sayings" of children.

Doctor—or shall we say Colonel—Ayres dodged the use of slides and brought his own charts. This was a wise precaution on his part, for we were having trouble with our lantern. These charts he comfortably set up and proceeded to offer an admirable paper on "The Comparative Effectiveness of State School Systems." He began by going back to 1750 at which time an Italian article was published on the high cost of living. In this article index numbers were used practically as they are now employed in economics. The speaker discussed the meaning and use of the index number and showed how it could be applied in education. He made the application on eight items to the schools of each of the states. These states were ranked on the basis of

each of the eight items according to their index numbers. Comparisons were also made with Dr. Ayres' previous study of conditions in 1910.

Last year, our association authorized a committee on standards. The president, however, seemed to have been particularly successful in selecting men who were about to retire from research work and from membership in the organization. The committee had no opportunity, therefore, to work out a real committee report. Mr. Courtis, however, as chairman, determined to have a report, so he wrote one himself. This he presented in part as the first number on the afternoon program. While certain recommendations were made for the guidance of the association and through it for the guidance of research workers in general, the paper consisted largely of the presentation of evidence to show that we really do not know exactly what our tests measure. As an indication of the analysis to which our testing instruments must be subjected, the paper was important.

Dr. Thorndike was to have presented a paper on "Intelligence Tests for College Entrants." He was, however, unable to attend—and again we have the "flu" to thank for his inability to respond. Dr. Thorndike sent his paper but, owing to the shortened program consequent upon the delay in getting started, the paper was not read. It will be published in an early number of the *JOURNAL OF EDUCATIONAL RESEARCH*.

Mr. Dickson—Dickson all the way from the Coast—gave a report on the use of intelligence tests in the classification of eighth and ninth-grade pupils. He demonstrated the reliability of these tests as administrative devices for the homogeneous grouping of pupils for educational guidance, and for providing specialized instruction for superior children.

Mrs. Fleming of the Wisconsin State Department presented her paper on "Reading Ability and Intelligence." Mrs. Fleming had results on the use of the Thorndike reading test in Wisconsin and Virginia. From the Wisconsin results she is tempted to conclude that there is a close correlation between reading ability and intelligence but the Virginia results seemed contrary. While not entirely committing herself, she seemed to feel that such correlation did exist wherever regular school attendance and emphasis upon reading were generally operative.

Dean Kelly presented a study of "Three Methods of Teaching the Fundamentals of Arithmetic." The methods were by use of (1) Courtis Practice Tests; (2) Studebaker Practice Tests; (3) Usual teacher-devised drill material. The results were definitely favorable to scientifically devised practice material, the "Courtis" getting, in general, the greatest gains for classes but the "Studebaker" evidencing greater corrective effect.

Mr. Burgess, one of our new members in the Russell Sage Foundation, presented "A Trend Analysis of Teachers' Salaries." By a study of the wages of skilled and unskilled laborers and of rural and city teachers during the past century and by converting the wages into index numbers, he showed the following facts: (1) wages of one kind follow the trend of wages of another kind much more closely than they do the cost of living, (2) wages of skilled and unskilled labor respond more quickly in economic crises than do teachers' salaries, (3) rural teachers' salaries and wages of unskilled labor tend to be

nearly equal except in times of crisis, (4) city teachers' wages tend to follow the same general curves as those of rural teachers and of skilled and unskilled laborers; but they lie between the wages of the last named groups, (5) times of high prices show a decided drop in the index of teachers' salaries.

Mr. Richardson showed some results attained by "The Campaign Method in Elementary Education" with special reference to arithmetic work in the four fundamental operations with whole numbers. The method consists of taking individual results directly to each child and interesting him in making a record for himself. Practice examples were furnished the pupils. Although the campaign was carried on for a few weeks only, results showed large and permanent gains.

The Fifth Annual Banquet and the evening program were the climax of the week to the members of our association. More than forty members and guests were present. Serious thought and kindly humor characterized the evening to the delight of all.

Following the practice of the past, the secretary presented one name for election to honorary membership. The association honored itself this year by electing Dean Wm. F. Russell of the College of Education, State University of Iowa. Dean Russell responded to the election with a striking allegory, portraying the evolution of the school superintendent in terms of the *chauffeur*.

Mr. Packer presented very briefly the need of a definite administration of supervisory work through a department of research. Such a plan is now in operation at Detroit. By this means, he said, much supervision which is ordinarily wasted, may be applied where it is most needed. His plan contemplates the fuller recognition of the building principal as a supervisor.

Dr. Rugg spoke on the necessity of scientifically devised text books as the solution of much of the present school inefficiency—a solution which can be applied even before we have an adequately trained body of teachers.

The first annual presidential address was given, summarizing the progress of the past year in the fields of particular interest to our membership. The magnitude of this progress was surprising to most of us.

At the business meeting, the following officers were elected for 1920-1921: President, M. E. Haggerty, Director of Bureau of Co-operative Research, University of Minnesota; Vice-president, Virgil E. Dickson, Director of Research and Guidance, Oakland and Berkeley, California; Secretary-Treasurer, E. J. Ashbaugh, Director of Bureau of Educational Service, State University of Iowa, Iowa City, Iowa.

Much of the business so far as discussion was concerned had been conducted at the closed session on Monday. Accordingly matters were disposed of quickly. In addition to the election of officers, the president was authorized to appoint a committee with power to act in the matter of affiliation with the National Research Council. He was also requested to appoint a permanent Standardization Committee and a committee to decide the day of the convention week on which our closed session shall hereafter be held. Dues were raised from \$2.00 to \$3.00 beginning with 1921.

Journal of Educational Research

Editorial Office, Urbana, Illinois

Editor-in-chief B. R. Buckingham, University of Illinois

VOLUME I

MAY, 1920

NUMBER 5

INTELLIGENCE EXAMINATIONS FOR COLLEGE ENTRANCE¹

EDWARD L. THORNDIKE

Teachers College, Columbia University

For an ideal examination of the intelligence of candidates for college entrance we might set the following specifications:

- I. *Significance*.—The score should correlate as closely with future achievement in college as is possible. This maximum possible correlation will not be 1.00, since achievement in college is due in part to health, to freedom from personal worries, and to various moral qualities as well as to intellect. Also, the magnitude of the correlation coefficient will depend on the range of the intellect of candidates, being smaller as that range is restricted. If all the eighteen-year-olds in the country were educated for college, tested, and given a trial in college, we might perhaps expect a maximum correlation as high as 0.75 to 0.85. Within the restricted range of those who complete a high-school course and actually become candidates, we may expect as a maximum 0.55 to 0.65, possibly more. A correlation above 0.50 is probably an improvement over the attainment of standard systems of accrediting high schools or of entrance to college by examination in school subjects.
- II. *Continued significance*.—The correlation should maintain itself year after year after the nature of the intelligence examination is known in spite of whatever coaching candidates undertake. This implies, in theory, that such

¹ A paper prepared for presentation before the National Association of Directors of Educational Research at its meeting at Cleveland, Ohio, February 26, 1920.

special coaching shall give training that is nearly as useful for future college achievement as for success in the test, or that the tendency to undertake, and the ability to profit by such special coaching shall be very closely correlated with ability to achieve in college. In practice, it means that the examinations year after year should be totally different in content until a great many of them are available; and should differ in form where there is danger that adaptation to the form of the examination may enable a dull boy to secure a good score.

III. *Comparability of standards and results.*—(a) The examinations given year after year should be closely equal in difficulty; (b) the scoring should be by standard keys so that the same paper scored by different examiners will receive substantially the same score; (c) the administration of the examination should be simple and independent of the personality of the examiner.

IV. *Probable error of the examination itself.*—For the examination's main purpose, the probable error is already taken account of in the correlations with college success. For that purpose, in fact, a large probable error would be encouraging, for it would mean that by repeating the examination, using a different form, the correlation with college success could be notably raised. For certain special purposes such as comparing individuals or groups, and for the justification of shortened forms of the examination, the smaller the probable error the better.

A half-year's experience with the Thorndike Intelligence Examination for High School Graduates may increase our confidence that these specifications may be met to a rather close approximation. This examination was devised at the request of the faculty of Columbia College, where it is an optional means of entrance for boys suitably recommended. Three forms of it have been used to date by Columbia, one in June 1919, one in September 1919, and one in February 1920. About seven hundred candidates have taken it, though over half of these were already legally admitted by the content examinations of Columbia University or the College Entrance Examination Board or the New York Regents. It has also been used for various purposes by a number of other institutions.

SIGNIFICANCE

Four years must elapse before an exact accounting on this point can be made, but the following facts may be noted:

Of eleven boys at Columbia, reported to the dean's office for inability to do college work in the early weeks of the year, all had notably low scores in the intelligence examination. Of a score or more so reported as a result of the mid-term records, all but two had low scores. The defective college work of these two was by common consent not due to intellectual defect.

In the case of a group of eighty odd candidates in another college, the correlation between the score in the Thorndike examination and the first term's college record was over 0.65. The highest 40 percent in the examination showed 74 percent of high grades in college work and 26 percent of low grades. The lowest 60 percent in the examination showed 27 percent of high grades and 73 percent of low grades. Of the 16 percent receiving the lowest grades in the examination, none received any college grade above C and 86 percent failed in five hours or more of their college work.

In a third institution, where the entire student body was tested, the members of the three upper classes showed no scores below 60 in the examination, whereas a third of the freshmen (admitted by certificate) scored below 60. The distributions given in the accompanying table (Table I) give a correlation of approximately 0.55 between test score and continuance beyond the freshman year. The returns received from the first and third of these institutions are a random selection from trials of the examination. Those from the second institution were reported to the writer. They may have been reported because of the positive evidence of the success of the examination, and so may be unduly favorable.

CONTINUED SIGNIFICANCE

If we take the best of the test material that has demonstrated high correlations, such as selected series of opposites, mixed relations, and cleverly devised problems and completions, and give it without fore-exercise, we can undoubtedly get high correlations, but we may be unable to maintain them.

Our opposites, relations, and problems of later years will probably not be so ingenious as those which we chose first. Moreover, part of the first high correlation may have been due to the testing

TABLE I. THE SCORES, IN PERCENTAGES, OF STUDENTS IN
A WESTERN PROFESSIONAL SCHOOL IN THE THORNDIKE
INTELLIGENCE EXAMINATION FOR HIGH
SCHOOL GRADUATES

(This institution admits on certificate from accredited schools but finds that a considerable percentage of the students so admitted do not satisfy its standards, the elimination being rather large. The figures below show that apparently an ability denoted by a score of 60 or more in the Thorndike examination is required for meeting the standards of this institution. None of the sophomores, juniors, and seniors fall below 60, whereas about one-third of the freshmen do.)

| Score in Thorndike Examination | [Freshmen Percent | Sophomore Junior Senior Percent |
|--------------------------------------|-----------------------|--|
| 30 | 2.6 | |
| 35 | 2.6 | |
| 40 | 12.9 | |
| 45 | 2.6 | |
| 50 | 2.6 | |
| 55 | 10.3 | |
| 60 | 15.4 | 5.5 |
| 65 | 10.3 | 16.3 |
| 70 | 12.9 | 10.9 |
| 75 | 5.1 | 8.1 |
| 80 | 7.7 | 19.0 |
| 85 | 10.3 | 10.9 |
| 90 | 7.7 | 8.1 |
| 95 | | 10.9 |
| 100 | | 8.1 |
| 105 | | 0 |
| 110 | | 0 |
| 115 | | 2.7 |

of ability to understand quickly the nature of the tasks. When the nature of the examination becomes a matter of public knowledge to candidates, the best features of it may be lost. It will be specially disastrous if some candidates do and others do not know about it.

Consequently, thorough precautions must be taken so that the examination will be as effective for twenty years as it is in its first trials. Our precautions are:

1. All test material is made up in quantities sufficient to provide for the entire series of examinations into which it enters, and is distributed without prejudice among them.² The selection for September 1929 will be nowise inferior to the selection for September 1919.

2. Three-fourths of the candidates' time is spent on paragraph reading, completing sentences, tests of information, arithmetical and algebraic computations and problem solving, common-sense questions, and following directions. The nature of these tasks is understood in a second or two by everybody of the intelligence level of a high-school graduate. In examining hundreds of papers, I have never seen a case of misunderstanding a single one of these tasks.

3. With two exceptions,³ fore-exercise is given for the other tasks, and two trials at each task are allowed. They include such tests as the disarranged sentences, opposites, number-series completions, verbal analogies, pictorial analogies, geometrical relations, and absurdities. The nature of these is, in general, such that any unfair advantage to a boy whom some unscrupulous tutor might train in test-taking would be slight. Experiments are planned to measure this gain exactly. If it is appreciable, the examination will be extended to include "traps" to penalize such improper preparation.

4. The content of the examination is entirely different in each issue. The form of the examination in the case of those tests where an unfair advantage from coaching is a possibility, will change every two or three years.

5. Finally, no new issue will be used regularly until it has been tried out sufficiently to insure its giving a prophecy substantially equivalent to the previous issues of the examination.

COMPARABILITY OF STANDARDS AND RESULTS

Since each element of the examination is selected at random as to difficulty, the total examination, comprising over twenty different tests, will in any one issue of the same general type, be closely equal in difficulty to any other (with the exception of the

² This is not true of the June 1919 series, which had to be arranged in part before the total material could be planned.

³ In the later series it is planned to provide fore-exercise for these, if any evidence appears that it is needed to equalize matters for those who have and those who have not previous acquaintance with the examination.

June, 1919 series which, for the reason given above, could not be fully standardized and is in part somewhat harder than the others). Each issue is further checked in this regard by actual trial. For example:

Eighty college students took Forms A and F, Part I, of the examination in that order. Eighty others took Forms A and G in that order. Letting a =score in A, f =score in F, and so on, we find that Median of $(f-a)$ =Median of $(g-a)+1$. That is, Form G is 1 point, or less than 1%, harder than Form F.

Thirty-six college students took Forms A and H in that order. Median of $(h-a)$ =Median of $(f-a)+1\frac{1}{2}$. That is, Form F is $1\frac{1}{2}$ points, or a little over 1%, harder than Form H.

One hundred fifty-four college students took Forms A and J in that order. Median of $(j-a)$ =Median of $(f-a)+2$. That is, Form F is 2 points, or about $1\frac{2}{3}\%$ harder than Form J.

Thus $f=g+1=h-1\frac{1}{2}=j-2$

Part II Form B and Part II Form C were each given to 85 college students, the two groups of 85 being random halves of a group of 170. The two groups were tested at the same time in the same room, by the same examiner, all instructions being identical. The median score for Form B was 81.5; for Form C, 79.0 or a difference of 2.5. The 25-percentile score for Form B was 58.3; for Form C 57.5 or a difference of -0.8 . The 75-percentile score for Form B was 102.7; for Form C 101.9 or a difference of -0.8 . In general, the two distributions are closely similar.

We may, therefore, conclude that Part II C is probably a trifle harder than Part II B (about $1\frac{1}{2}$ percent).⁴

Thirty-three college students took both III B and III C, sixteen having one form first and seventeen having the other form first. The median of $(c-b)=1\frac{1}{2}$ points or about $2\frac{1}{2}$ percent.

Two examinations made up respectively of:

| | | | |
|-----|-----|------|-----------|
| I G | I J | II B | and III B |
| I F | I H | II C | and III C |

would thus be almost exactly equal in difficulty for a college group.

Twenty-three adults took both II B and II D, twelve having Form B first and eleven having Form D first. The median of $(b-d)=4$. That is, Form D is 4 points harder than Form B.

⁴ Test 7 was omitted in both cases, but we know from other evidence that II B 7 and II C 7 are very closely similar in difficulty.

So, when II D is used in place of II B or II C, the other parts of the examination should be chosen to be a trifle easier than I G, I J, III B, or I F, I H, III C.

The difference in difficulty between any two issues can thus be known and allowed for by transmuting the score, or, if it is at all large, it can be reduced by alterations in the examination and re-checking. This process of equalization can be applied to the 10-percentile or 20-percentile or any other ability levels as well as to the general tendency of the group.

This checking is extremely laborious and probably unnecessary, since differences of 3 and 4 or even of 5 and 6 percent in the difficulty of different issues are of little practical consequence in any uses to which the examination is likely to be put. However, their reduction to small amounts and an exact knowledge of these amounts make an instrument available for very precise comparisons.

The scoring of all save the reading and completion tests is, with very slight exceptions, done by absolute keys with no requirement of judgment on the part of the scorer. The scoring of the reading and completion tests requires some judgment in the use of the keys, but the correlation between the ratings given by two scorers of the ability of a college instructor or assistant is very high. The correlation between two such scorers is 0.98. The probable difference between two such scorers was found to be only five points out of a total score which ranges from 7 to 96 among the actual candidates whose papers were scored. These scorers used only the stock keys supplied with the tests. By having one person score one test of the six, another score another, and so on, the probable error for the total result of one such group of six would obviously be unappreciable. This is the practice recommended where the examination is used with groups of any considerable size.

The administration of the examination consists simply of giving out the blanks, and instructing the candidates to go ahead at certain times even if they have not finished the work to date. In general, a candidate does test after test without awaiting instructions.

IMPROVEMENT OF THE EXAMINATION

As soon as adequate records of success in college are available, the correlation of each element of the examination therewith will

be computed together with the partial correlations where these are instructive. In the meantime, data are being gathered concerning the value of different available tests as shown by their correlations with a composite score obtained in seven or eight hours of work by adults of the general intelligence level of college candidates. Some of the data so far obtained are of value to anybody who is working with group tests for adult intelligence, and are presented here. Lack of space forbids adequate description of the tests. The following definitions apply to captions appearing in Tables II and III.

1. "Part I (first half)" means the first 30 minutes work of the standard examination as now given.
2. "Part II" means Part II of the standard examination.
3. "Part III" means Part III of the standard examination.
4. "New Part I (first half)" means a new form of Part I to go into effect in 1922.
5. "Information" means a test of the type used in the Army Alpha, but harder.
6. "Name part" means a test in writing a word that means a part of the thing named by the given word, and also begins with a certain letter.
7. "Name opposite" means a test in writing a word that means the opposite of the given word, and also begins with a certain letter.
8. "Name thing of which is part" means a test in writing a word meaning something of which the thing named by the given word is a part. The word written must also begin with a certain letter.
9. "Name thing made of," "synonyms" and "species of genus" are like the above. The word written has to fulfill the requirement of being in the stated relation to the given word and also of beginning with a certain letter.
10. "Sum, difference, product" means a test in giving two numbers which are defined by certain relations (e.g. two numbers whose product is 1 and whose sum is 2).

TABLE II. CORRELATIONS OF VARIOUS FORMS OF GROUP TESTS
OF ADULT INTELLIGENCE LEVEL WITH A COMPOSITE
SCORE FOR SEVEN HOURS OF TEST WORK, IN
THE CASE OF 33 COLLEGE STUDENTS

| Test | Time in Minutes | Correlation |
|----------------------------------|--------------------|-------------|
| Part I (first half)..... | 30 | 0.75 |
| Part II..... | 60 | 0.85 |
| Part III..... | 40 | 0.65 |
| Part III..... | 120 | 0.81 |
| New Part I (first half)..... | 30 | 0.84 |
| Information..... | 15 | 0.62 |
| Name part..... | 8 | 0.40 |
| Name opposite..... | 12 | 0.80 |
| Name thing of which is part..... | 13 | 0.72 |
| Name thing made of..... | 13½ | 0.61 |
| Synonyms..... | 16 | 0.66 |
| Species of genus..... | 12½ | 0.61 |
| Sum, difference, product..... | 20 | 0.83 |
| Total Time..... | 390 | |

TABLE III. INTERCORRELATIONS OF VARIOUS FORMS OF GROUP
TESTS FOR ADULT INTELLIGENCE LEVELS
33 COLLEGE STUDENTS

| | Name Part | Name Opposite | Thing of Which Is Part | Thing Made of | Synonyms | Species of Genus |
|--------------------------|--------------|------------------|---------------------------------|---------------------|----------|------------------------|
| Information..... | 0.03 | 0.56 | 0.53 | 0.54 | 0.48 | 0.52 |
| Name part..... | | 0.25 | 0.27 | 0.30 | 0.30 | 0.33 |
| Name opposite..... | | | 0.61 | 0.46 | 0.83 | 0.49 |
| Thing of which part..... | | | | 0.49 | 0.64 | 0.39 |
| Thing made of..... | | | | | 0.49 | 0.55 |
| Synonyms..... | | | | | | 0.21 |

Facts like those shown in Tables II and III are being collected from individuals of the intelligence level of college students as fast as time and facilities permit, so that no test need ever be used in the examination until its value has been demonstrated.

A NEW METHOD OF RATING TEACHERS

WILLIAM L. CONNOR¹

Superintendent of Schools, Republic, Michigan

It is the teacher's task to make changes for the better in the abilities, habits, and attitudes of boys and girls. Her efficiency can be evaluated fairly only in terms of her success at this task. In other words, if a teacher is rated at all, she should be rated, not by the clothes she wears, or the method she chooses, but by the results she secures.

The rating card here proposed, together with the lists of "concrete acts" and suggestions for use, constitutes a scheme for rating teachers according to the results they secure with boys and girls. It is not a perfect plan; but it has proved more satisfactory to teachers than the usual type of rating scheme in which personal qualities, training, discipline, attention to equipment, etc., take equal rank with the results secured. The studies reported in this paper were made by persons in both supervisory and teaching positions, who were trying to make a "promotion by merit" scheme work. They believe they have evolved a way of rating the efficiency of teaching which is better than any scheme in common use, but realize it is still vastly short of the "dream of the scientist in education" of which Coffman speaks:² a scale for measuring the efficiency of teaching.

The stimulus to the evolution of the new rating scheme was the adoption by the board of education of Republic, in the fall of 1916, of a resolution creating a uniform basal salary of \$550.00 a year for the grades and \$750.00 for the high school, and also empowering the superintendent to rate the efficiency of each teacher as a basis for determining additional compensation ranging from \$25.00 to \$75.00 a year. These figures were later increased to \$650.00 and \$850.00 per year and the range of the bonuses from \$50.00 to \$150.00.

¹ The author appreciates and desires to acknowledge the assistance given by teachers, and others including the editors, of the JOURNAL OF EDUCATIONAL RESEARCH in the preparation of this study.

² Coffman, L. D. "The rating of teachers in service," *School Review Monograph*, No. 5, p. 14 ff., 1914.

Only a short time before, the scheme of paying labor according to the product had been adopted in the mines about Republic. The "bonus for merit" plan was thus readily comprehended by the board of education and adopted as a means of meeting the teacher situation. The details of the plan for distributing the additional compensation for merit were left to the superintendent to devise. As at first proposed and adopted, it was in general as follows: Teachers were to be rated by the superintendent and divided into five classes, namely, inferior, below average, average, above average, and superior. Teachers in the two lowest classes were not to receive any extra compensation, the others \$25.00, \$50.00, and \$75.00 in order of merit. Later these bonuses were increased to \$50.00, \$100.00, and \$150.00 a year.

The first simple rating scheme was composed of material from the studies of Ruediger and Strayer,³ Elliott,⁴ Boyce,⁵ and others, and was used for the first time in January, 1917. Cubberley had only recently written that some such plan was the most important means of stimulating and rewarding professional growth and efficiency. At the same time, he had warned that it was the hardest plan to put into successful operation. Other school men whose opinions the writer valued asked him, some jocosely, some seriously, if he expected to weather the storm raised by the first ratings. It was, then, with mingled hopes and misgivings that the plan was presented to teachers for their consideration and criticism.

During the period when the details of the plan were being worked out, the teachers were being invited to examine the plan and were generally brought to approve its logic. At the same time they were warned that the rating scheme proposed for use was very imperfect and that the superintendent might make mistakes however just he might try to be. The immediate increase of \$50.00 a year to over half the teachers, together with the prospective increases for merit, tended to stay criticism for a while. This time was used in getting the teachers ready to attack the

³ Ruediger, W. C., and Strayer, G. D. "The qualities of merit in teachers," *Journal of Educational Psychology*, 1:272-78, May, 1910.

⁴ Elliott, E. C. "How shall the merit of teachers be tested and recorded," *Educational Administration and Supervision*, 1:291-99, May, 1915.

⁵ Boyce, A. C. "Methods for measuring teachers' efficiency," *Fourteenth Yearbook of the National Society for the Study of Education*, 1915, Part II, p. 10.

rating scheme, instead of the superintendent or the logic of the plan, when in due time they should become critical. Little by little they were brought to feel that they were about to cooperate in doing two of the hardest and most worthwhile things in education; namely, working out a satisfactory scheme for rating the quality of teaching, and using the scheme in determining a part of the pay of teachers.

The first formal ratings were passed out in January, 1917. The superintendent found no superior teachers and only one inferior teacher. About one-fourth of the force was rated above average, one-half average, and one-fourth below average and inferior together. Bonuses were paid for the first half of the year. Three-fourths of the teachers received checks for \$12.50 or \$25.00 with their first ratings. The expected criticisms appeared at once, but the lines for it to follow had been well enough drawn to confine it largely to really helpful suggestions for revising the rating scheme.

After the first general rating, the teachers became intelligently critical and cooperated actively in revising the method of rating used. Various schemes were tried which need not be described here. One involved an attempt to make standard definitions of merit.

A small group of superintendents and an equal number each of principals, supervisors, and very capable teachers were asked to choose, from among the teachers of their acquaintance, one representative of each of the five classes proposed here, namely, inferior, below average, average, above average, superior; and, then, to describe at work the teachers thus chosen to represent the several classes, or merit-groups, using the "First Rating Scheme Devised on a Cooperative Basis, 1917" as an outline for the description (Figure 1). Quite naturally the scheme fell through. We could not persuade our friends literally to write a book for us in their spare time. Gradually, the conviction grew in the minds of the teachers and the superintendent that the scheme gave too much credit to items only remotely related to the real business of the school, namely, making changes for the better in the behavior of boys and girls. Here I use *behavior* in the sense of all significant mental and physical responses to all situations.

FIGURE 1. FIRST RATING SCHEME DEvised ON A COOPERATIVE BASIS, 1917

Teacher.....Grade or subject.....
 School.....Address.....Date.....19.....

The Teacher and Her Work

- | | | |
|------|--|-----------|
| I. | The teacher..... | 5 4 3 2 1 |
| | A. Personal appearance..... | 5 4 3 2 1 |
| | B. Character..... | 5 4 3 2 1 |
| | C. Education..... | 5 4 3 2 1 |
| | D. Social efficiency..... | 5 4 3 2 1 |
| II. | The government of the school..... | 5 4 3 2 1 |
| | A. Assignments..... | 5 4 3 2 1 |
| | B. Morale in the despatch of assignments..... | 5 4 3 2 1 |
| | C. Freedom to initiate social activities..... | 5 4 3 2 1 |
| | D. Elimination of anti-social conduct..... | 5 4 3 2 1 |
| III. | Instruction as providing for educative activities..... | 5 4 3 2 1 |
| | A. Thinking..... | 5 4 3 2 1 |
| | B. Emotional reaction..... | 5 4 3 2 1 |
| | C. Acquisition of knowledge and skill..... | 5 4 3 2 1 |
| | *D. Deportment..... | 5 4 3 2 1 |
| IV. | The teacher's attention to..... | 5 4 3 2 1 |
| | A. Health and comfort of pupils..... | 5 4 3 2 1 |
| | B. Play and social life of pupils..... | 5 4 3 2 1 |
| | C. Grounds, buildings, equipment & supplies..... | 5 4 3 2 1 |
| | *D. Professional ethics..... | 5 4 3 2 1 |

NOTE: Qualities are designated thus:—superior, 1; above average, 2; average, 3; below average, 4; inferior, 5.

* These two items were added during 1917–1918.

While in this state of mind, the author listened to an address given before the Atlantic City meeting of the National Association of Directors of Educational Research in which Assistant Superintendent Allison of Chicago stated, by way of illustrating a point about the Boyce score card, that it was his belief that the correlation between personal appearance and teaching ability was zero or actually negative.⁶ Another man at the same meeting wittily remarked: "The homely girls make the principles work, while the pretty ones work the principals—p-l-e-s and p-a-l-s, please note." Only recently Morton's⁷ study reveals much the

⁶ Bulletin No. 1 of the Board of Education of Chicago, November 23, 1917, contains much of the material used by Mr. Allison in this address.

⁷ Morton, R. L., "Qualities of merit in secondary teachers," *Educational Administration and Supervision*, 5:225–38, May-June, 1919.

same general state of affairs. He says in one place: "Evidently, then, in the judgment of the inspector, a desirable reaction on the part of the pupils cannot (certainly) be secured by a teacher whose personality and scholarship are ideal, whose methods are above reproach, and who teaches in a room where the physical factors are as they should be." In the judgment of the writer, a perfectly groomed teacher of perfect physical proportions working with perfectly graded children in a perfect physical environment *may* and often *does*, use perfectly good methods in a perfectly futile way.

This rather sweeping conclusion was pretty generally reached by teachers in Republic in March, 1918, and it was very soon afterward that they adopted certain guiding principles for their further study of rating teachers. These principles, briefly stated, are as follows:

1. Teaching, not teachers, should be rated.
2. Teaching is measurable in terms of the results only.
3. Results may be classified as inferior, below average, average, above average, or superior, only with reference to the amount of improvement made and the ability of the particular group of boys and girls to improve.

Once these principles are accepted, four general problems present themselves for solution. They are as follows:

1. Determination of a few general headings under which all the items of behavior the school desires to affect favorably through the agency of the classroom teacher can be grouped.

2. Evaluation of these headings, or groups of items of behavior, in the desired result.

3. Choice, from existing scales, of those which measure progress along the lines desired, and the creation of new scales needed, or other means of making rough estimates of progress along lines where scales are yet far off.

4. Determination of how much progress any particular group assigned to any classroom teacher ought to make.

The scheme for rendering standards fairly intelligible, which was finally decided upon as the only feasible one for Republic, was this:

1. All items on the rating card were defined in general terms, with reference to the dictionary.

2. Items were redefined in terms of the standards thought to be desirable in education.

3. As¹ many concrete acts and conditions of teachers, teaching, pupils, schoolrooms, etc., as could be collected by the superintendent and the cooperating teachers were typewritten and made available for use in supervision. This material grew rapidly from a few pages to more than a hundred.

The seven elements or standards by which teachers were to be rated were mimeographed on separate slips of paper and distributed, with appropriate explanations and directions to those taking part in the experiment.

STANDARDS

I. THINKING

General definition: Mental activity in the course of which the mind locates a difficulty, and in the process of solving the difficulty arrives at a conclusion which is, to it, a new meaning, belief, act or way of thinking, or any combination of these. *Definition of standard:* Thinking by pupils in study or recitation, and test of the quality and amount agreed upon and set down in the course of study or other practice standards.

II. EMOTIONAL REACTION

General definition: Feelings, and actions growing out of feelings. *Definition of standard:* Appreciation resulting in immediate pleasure, and emotion leading to action which is of individual and social value.

III. KNOWLEDGE AND SKILL

General definition: Information and practical efficiency in doing. *Definition of standard:* Classroom practices and results as to knowledge and skill as set down in the course of study and other practice standards. (A course of study in which standards of procedure and results as to knowledge and skill required are not determined by objective measurement of both needs and results, in so far as that is possible today, is hardly to be defended.)

IV. MORALE IN DESPATCH OF WORK

General definition: Confidence, courage, loyalty, self-reliance, promptness, and persistence with which work is performed. *Definition of standard:* Confident, courageous, loyal, self-reliant, prompt, and persistent spirit of pupils at work.

V. INITIATIVE IN SOCIALLY SIGNIFICANT SITUATIONS

General definition: Ability and will to propose or choose, lay out, and carry forward work furthering the purpose of the individual and the group. *Definition of standard:* Pupils proposing or choosing, organizing, and carrying forward work furthering the educative purpose of school groups.

VI. ETHICAL SELF-CONTROL IN SITUATIONS SOCIALLY SIGNIFICANT

General definition: Inhibition of actions which would interfere with carrying forward the purposes of social groups. *Definition of standard:*

Inhibition in the classroom, and other group-units of the school, of conduct interfering with carrying forward educative activity.

VII. DEPORTMENT

General definition: The manner, bearing, or style of a person in doing things. *Definition of standard:* Easy, natural, graceful, and courteous use of correct form in doing things.

INSTRUCTIONS ACCOMPANYING THE STANDARDS

THINKING STRAIGHT

All of you who arrange these standards want to express your best judgment of the relative values of the standards. In order to do this, avoid thinking of the *time order* in which you may have been taught to believe these standards are acquired by boys and girls. Think rather of the concrete acts making up each standard as being considerably different at different ages. More concretely, think of the group of boys and girls you teach as possessed, and not possessed, alternately, of the attitudes, abilities, and habits implied in each standard, and ask yourself: "In what *order of importance* would I place the *possession* of these standards by my boys and girls? How would I *emphasize* their acquisition by the pupils under my charge if I felt a deep responsibility to each individual in my group, and to society as a whole, and knew my efforts this year were to be the last conscious educative efforts to be expended on a majority of these boys and girls for the rest of their lives?" But do not rate any proposed standard low because you do not now know how to get pupils to acquire it. Rate the standard according to its importance or desirability, and face squarely the issue of finding new subject-matter or methods by which the standard may be realized. In this concrete fashion, test and re-test your arrangement of the standards.

FINAL DIRECTIONS FOR ARRANGING THE STANDARDS

1. Read the seven standards critically, trying to evaluate them.
2. Spread the standards out before you so that you can get what they contain in perspective.
3. Arrange the slips on which standards are printed in the order of the importance of the standards in the results of teaching, beginning with the most important.
4. If, in your judgment, the least important standard is so low as to approach zero, mark it zero.
5. Pin slips together in the order in which you rank the standards and return them in that order.

One hundred and thirty-two persons, all of whom were engaged in school work, assisted in the study by arranging these standards. Fifty of the judges worked under the personal direction of the author, and 45 others under the direction of one or another of the first fifty. The remainder of the judges received the material by mail. Seven of the replies were rejected because it was apparent that the persons making them had not followed directions. The returns from 125 judges were tabulated.

A few less than a hundred judges may be grouped into four divisions as follows: A group of graduate students in a certain university, a group of local teachers, a group of teachers and teachers-in-training in a city normal school, and a group made up of elementary and junior high-school teachers working under one principal in a large city. The remainder of the judgments were made by principals and superintendents in various Michigan towns and cities.

The data derived from the 125 judgments are as follows:

TABLE I. FREQUENCY WITH WHICH EACH RANK
WAS GIVEN EACH ITEM

| Item No. ^a | Rank ^b | | | | | | | |
|-----------------------|-------------------|----|----|----|----|----|-------|-------|
| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| I..... | | | | | 11 | 7 | 28 | 79 |
| III..... | | 9 | 9 | 13 | 23 | 23 | 35 | 13 |
| V..... | | 3 | 12 | 23 | 19 | 31 | 23 | 14 |
| IV..... | | 6 | 10 | 27 | 29 | 28 | 16 | 9 |
| II..... | | 12 | 23 | 20 | 28 | 24 | 12 | 6 |
| VI..... | | 18 | 43 | 32 | 12 | 4 | 12 | 4 |
| VII..... | 1 | 76 | 29 | 9 | 3 | 7 | | |

^a The higher the rank, the larger the number indicating the rank. Items of such small value as to be insignificant, in the opinion of the judges, are scored 0.

^b The number of items in the original material are preserved for reference.

Table I should be read as follows: Seventy-nine judges thought Item No. I most important, 28 thought it second in importance, seven thought it third in importance, and eleven thought it fourth in importance. None thought it less than fourth in importance.

The returns for the different groups have also been tabulated separately. From these tabulations the following facts appear:

1. The order of importance of the items is the same in the case of each of the groups of judges.

2. The general form of each distribution for each item is the same for each of the groups of judges.

3. The judges who worked under the personal direction of the author recognized the real differences in relative position as established by the whole group a little more frequently than did the judges to whom the material was sent by mail.

It is probably safe to conclude, therefore, that the items do in reality stand in the order of importance here assigned. On the other hand, the importance of this order may be great or small according as the absolute value assigned to the least important item in the list is great or small.

From Table I, the values of the various items in terms of probable error units have been determined and corrected into point scales of 100 and 1,000 points.

TABLE II. THE VALUES OF THE ITEMS

| Value of Intervals | | Cumulative or Total Values | | Percent Each Value Is of Total Value | ADJUSTED POINTS | |
|--------------------|-----|----------------------------|------|--------------------------------------|-----------------|---------------|
| | | | | | Basis of 100 | Basis of 1000 |
| 0-VII | 3.0 | VII | 3.0 | 7.37 | 7 | 75 |
| VII-VI | 1.6 | VI | 4.6 | 11.30 | 11 | 115 |
| VI-II | 1.0 | II | 5.6 | 13.76 | 13 | 135 |
| II-IV | 0.3 | IV | 5.9 | 14.50 | 14 | 145 |
| IV-V | 0.4 | V | 6.3 | 15.48 | 16 | 155 |
| V-III | 0.2 | III | 6.5 | 15.97 | 17 | 160 |
| III-I | 2.3 | I | 8.8 | 21.62 | 22 | 215 |
| Totals | 8.8 | | 40.7 | 100.00 | 100 | 1,000 |

The value of the various items varies according to the value assigned the least important item in the list. For example, if Item No. VII were zero, as one judge would have it, Item No. I would be 5.8 units, or three and six-tenths times the value of Item No. VI, under the same condition. Whereas, under the value assigned to Item No. VII, Item No. I is only one and nine-tenths times the value of Item No. VI. On the other hand, if Item No. VII were assigned a value of 5.0 units, Item No. I would be only about two times the value of Item No. VII; whereas, under the value now assigned to Item No. VII, Item No. I is three times as important as Item No. VII. Without going any further into the need for, and difficulties involved in, trying to fix a value for one of the items with reference to zero, it may be said that three units were adopted as the value of the least important item, mainly, because the author was in constant contact

with a number of judges who were willing, for the most part, to allow a value that high for the item in question, but were unwilling that the value should be placed any higher. The reason, then, for assigning a value of three units to Item No. VII, *deportment*, was a practical one rather than a scientific one.

The rating scheme which follows is based upon the values given in Table II.

FIGURE 2. TENTATIVE PLAN FOR RATING A TEACHER ON THE PROGRESS SHOWN BY PUPILS

Teacher _____ Training _____
 Experience _____
 Grade _____ Subject _____ Date _____

Rough estimate of GENERAL MERIT: A B C D E (Circle thus: ⑥)

Analysis of results or progress by observation, examination, or application of standardized tests or scales, or by two or more of these methods:

| GENERAL MERIT..... | E | D | C | B | A |
|------------------------------|-----|---------|----------------------------|-------------|-------|
| I. Deportment..... | 1 | 2 | 3 4 5 | 6 | 7 |
| II. Ethical self-control.... | 1 | 2 3 | 4 5 6 7 8 | 9 10 | 11 |
| III. Emotional reaction.... | 1 | 2 3 4 | 5 6 7 8 9 | 10 11 12 | 13 |
| IV. Morale..... | 1 | 2 3 4 | 5 6 7 8 9 10 | 11 12 13 | 14 |
| V. Initiative..... | 1 | 2 3 4 5 | 6 7 8 9 10 11 | 12 13 14 15 | 16 |
| VI. Knowledge and skill... | 1 | 2 3 4 5 | 6 7 8 9 10 11 12 | 13 14 15 16 | 17 |
| VII. Thinking..... | 1 2 | 3 4 5 6 | 7 8 9 10 11 12 13 14 15 16 | 17 18 19 20 | 21 22 |

Summary: Numerical score _____ Equivalent: A B C D E (Circle thus ⑥)

Explanation: Numbers for each item in the groups A, B, C, D, and E give the following approximately correct numerical values for each group: A (superior) 93-100; B (above average) 72-92; C (average) 30-71; D (below average) 9-29; E (inferior) 0-8.

SUGGESTIONS FOR IMPROVEMENT:

(Criticisms interpreting the rating given usually cover several type-written pages, and contain references to the list of "concrete acts" appearing below, results of standard tests given to pupils, and the principles of education as expressed in standard books on education and psychology. *A brief summary of this material appears here.* Copies are kept for reference and comparison.)

 Superintendent

While still working with the typical mixed rating scheme prepared as the result of the first cooperative patchwork on rating plans, considerable progress was made in collecting concrete examples of each standard proposed. Enough was done to make it clear that the bulk of such material, if collected, would be enormous. While work was progressing on the evaluation of the seven general headings, some attempt was made to choose applicable material from the earlier study and restate it as a list of concrete acts, which, if observed in pupils, would be some evidence that the standards under which they were classified were being acquired. To this material were added statements concerning the standard tests and scales in use or proposed for use in Republic schools. Lastly, teachers contributed an enormous amount of material so detailed that it was thought necessary for present purposes to have a committee classify, arrange, and condense it into the limits of seven mimeographed pages.

The list of concrete acts which follows represents a condensed statement of the supporting material used with the proposed scheme for rating teachers on the basis of the results of their teaching. Any teacher, pupil, or supervisory officer, attempting to use this material will of necessity choose only those items applicable to the case in hand. Many teachers have prepared for themselves much more complete and detailed examples under each standard, but they fall under some of the headings given.

CONCRETE ACTS

STANDARD I. THINKING

1. Attains standard score for his grade in
 - a. Stone's Reasoning Tests (grades 3-8)
 - b. Composition (grades 4-12)
 - c. Trabue Language Scale (grades 2-12)
 - d. Thorndike's Visual Vocabulary Scales (grades 3-12)
 - e. Thorndike's Scale Alpha 2 for Measuring the Understanding of Sentences (grades 3-12)
 - f. Starch's Physics Test
 - g. Henmon's Latin Tests, *etc.*
2. Attains normal school grades in
 - a. School tests over minimum essentials in arithmetic
 - b. Teacher's or superintendent's test in other subjects
 - c. School tests in grammar (grades 5-8).
3. Expresses himself clearly in his own words.
4. Exhibits ready ability in reproducing in recitations the thinking that has been done before in mastering past work.

5. Interprets intelligently any complex directions for work.
6. Classifies and arranges material well.
7. Defines accurately, though sometimes awkwardly. (In the kindergarten by pointing and gesturing.)
8. Locates and defines difficulties in material suitable for his grade.
9. Proposes solutions to problems cautiously and tries them, selecting and rejecting material intelligently.
10. Reaches tentative conclusions and holds them tentatively while searching for more evidence.
11. Verifies conclusions carefully before considering them final.
12. Considers the bearings of conclusions carefully.
13. Accepts the true bearings of conclusions when they are pointed out.
14. Evaluates problems as well as the material used in the solution of problems.
15. Decides correctly when to depend on his own resources and when to seek help.

STANDARD II. KNOWLEDGE AND SKILL

1. Attains standard scores for his grade in
 - a. Courtis Standard Research Tests, Series B, in fundamentals of arithmetic (grades 4-8)
 - b. Woody's Arithmetic Scales (grades 2-8)
 - c. School tests in minimum essentials in arithmetic (grades 3-8)
 - d. School tests in minimum essentials in grammar (grades 5-8)
 - e. Spelling words from Ayres Spelling Scale concealed in sentences
 - f. Composition, technic (grades 2-12)
 - g. Copying familiar material (rate and quality) Ayres Handwriting Scale "Gettysburg Edition"
 - h. Stone's Reasoning Tests (grades 3-8)
 - i. Trabue's Language Scale (grades 2-12)
 - j. Starch's Physics Test
 - k. Thorndike's Visual Vocabulary Scale (grades 3-12)
 - l. Thorndike's Scale Alpha 2 for Measuring the Understanding of Sentences (grades 3-12)
 - m. Henmon's Latin Tests
 - n. Rugg and Clark's Standardized Tests in First-Year Algebra
2. Attains normal grades on teacher's and superintendent's tests in subjects not covered by standard tests.
3. Exhibits daily in the recitation a ready knowledge of facts and skill in fundamental processes of the subject in hand.

Note: Whatever value is assigned to items 2 and 3 should be apportioned about equally among the subjects for which the teacher is responsible, and for which satisfactory standard tests have not yet been devised.

STANDARD III. INITIATIVE IN SOCIALLY SIGNIFICANT SITUATIONS

1. Asks questions worth while for the whole group to answer.
2. Proposes problems essential to the progress of the group in the subject at hand.

3. Examines material in advance so as to be of service to the teacher and the class in laying out the work for the following day or week; i.e., assists actively in making the assignment.
4. Volunteers information, and makes helpful suggestions in the course of the recitation.
5. Seizes upon class problems for further reading or experiment.
6. Prepares to contribute intelligently, and offers promptly and eagerly to keep the topic, question, or problem going during the recitation.
Note: Many recitations make one think that teaching, like mining (or dentistry), is an *extractive industry*.
7. Calls attention to current literature, discoveries, inventions, or public affairs as related to topics or questions on hand.
8. Proposes for group decision the game to be played during the play or language periods.
9. Proposes changes in games and discusses the changes with teacher and pupils.
10. Is not idle when assigned work is finished, but finds useful work to do.
11. Discovers his weak points and practices to overcome them.
12. Does many useful things without being told.
13. Asks for needed information or suggestions when his own resources fail.

STANDARD IV. MORALE

1. Attends school regularly, and arrives promptly at the opening of each session.
2. Is cheerful, and agreeable when he cannot have his own way.
3. Moves to and from classes and about halls in a prompt and orderly manner.
4. Is loyal to teachers and the work they assign.
5. Is loyal to school officers and the rules they make and enforce.
6. Is loyal to the school and its undertakings.
7. Obeys all ordinary requests promptly and cheerfully, and in spirit as well as letter.
8. Does not talk back or sulk when corrected.
9. Does not fret or worry over school tasks, but seeks to understand them and does his best cheerfully.
10. Does not cry, whine, or tattle over little things.
11. Is not discouraged by defeat or failure.
12. Does hard or otherwise disagreeable work without expectation of praise.
13. Works in the confidence that he and his classmates can do what other boys and girls have done and are doing, and a little more.
14. Relies upon his own efforts in preparing a lesson.
15. Undertakes with courage work he knows to be difficult.
16. Finishes assigned lessons even if he is compelled to spend more time on them than some of his brighter classmates.
17. Begins work promptly and plunges into the heart of it.
18. Faces duties and responsibilities squarely—does not “sidestep” or “pass the buck.”
19. Does not hunt for and elect supposed “snap courses.”

20. Is conscious of some of the important habits which go to make up a worthy character, and strives to act so that these habits will be formed.

STANDARD V. EMOTIONAL REACTION

1. Appreciates standards of achievement in school subjects and wants to attain them.
2. Appreciates standards of achievement in school subjects and wants his class and his school to attain them.
3. Takes pride, but not foolish pride, in his class, school, city, state, or nation.
4. Takes pride in superior workmanship, and will not do slipshod work.
5. Takes pride in the character of his associates and chooses them so that he will not be ashamed of them.
6. Takes pride in cleanliness of person, schoolroom, premises, and streets.
7. Wants to serve his class, school, town, state, and nation.
8. Enjoys good music, literature, and art.
9. Likes innocent fun on every proper occasion.
10. Enjoys school exercises:—rhythm, story-games, free games, etc.
11. Gets a true emotional reaction out of insight into external truth either in science or in religion and morals.
12. Feels and expresses reverence for ties of friendship, family, and religion.
13. Does any desirable act because he wants to be helpful, and refrains from undesirable acts because he dislikes them.
14. Appreciates kindly motives in others and gets satisfaction out of expressing his appreciation.
15. Values the personality of others and his own as well.
16. In the classroom acts of service are popular.

STANDARD VI. ETHICAL SELF-CONTROL IN SITUATIONS SOCIALLY SIGNIFICANT

1. Obeys requests promptly and cheerfully.
2. Obeys the rules of the school.
3. Enters the room and leaves the room quietly and inconspicuously, especially during school hours.
4. Avoids coughing or spitting in public.
5. Does not intrude on another person's privacy or take his time unjustly.
6. Does not "delay the game."
7. Tells the truth without evasion.
8. Borrows little and returns borrowed articles promptly and in good condition.
9. Does not steal or even borrow without permission.
10. Keeps appointments.
11. Keeps promises made, and avoids rash promises.
12. Does not copy work or seek unfair help from others.
13. Will not permit copying or give unfair help to others.
14. Does not "bluff" in the recitation.
15. Does not bribe, trade, blackmail (threaten to tell) or use any other undue influence to have his own way in athletics or other school affairs.
16. Controls temper and avoids quarrels.
17. Does not allow himself to be bribed, intimidated, or influenced unduly in any way in school affairs.

18. Does not tattle, but will not allow the interests of the school to suffer or injustice to be done because of lack of information he could supply.
19. Does not tease or bully weaker persons nor allow others to do so without protest.
20. Cooperates in protecting school buildings, premises, and streets from litter such as fruit skins, paper, and marking.

STANDARD VII. DEPORTMENT

1. Sits well back in seat leaning forward from his hips, feet on floor, head erect, and forearms resting easily on desk when writing or working at desk.
2. Rises easily from seat, stands erect on both feet, free from desk, in reciting and sits again easily.
3. Walks or runs lightly in moving about classrooms and halls.
4. Speaks and reads clearly and distinctly so that every person in the room may hear.
5. Holds book or paper in correct position when reading orally.
6. Passes and collects materials quietly and easily.
7. Takes out and puts away work quickly, quietly, and neatly.
8. Puts on and takes off wraps quickly and quietly.
9. Does not stop work to stare at visitors.
10. Is courteous:
 - a. Is unassuming.
 - b. Avoids abrupt speech and coarse manners.
 - c. Laughs and talks quietly.
 - d. Does not rush or crowd in halls.
 - e. Allows others to enter doors ahead of him.
 - f. Keeps to the right.
 - g. Avoids passing between persons in conversation or in front of anyone.
 - h. Grants a favor or acknowledges one easily.
11. Does not put pencils or pen in the mouth, or fingers in mouth or nose.
12. Dresses neatly and is clean.
13. Keeps working materials in order on desk, and at the board.
14. Turns in neat work.
15. Responds, when addressed, pleasantly and in good form, answering questions, "Yes, Miss."; "No, Miss."; or giving information in clear sentences.

A careful reading of the above list of concrete acts will reveal that the means at hand for making scientific observations of the progress of pupils are scanty. It is still necessary, if one wishes to know whether pupils are making progress in forming most of the desirable habits and attitudes which they ought to form as the result of their school experience, to watch them at work and play in all the situations for which the school is responsible. Consequently, it would be foolish to assert that the most careful observer can detect more than the general trend of habit and

attitude formation outside of the limited field now occupied by the scales and tests. This is one of the strong points made by teachers against the author's use of the rating scheme in rating their work.

It does not seem at all impossible, however, that, should we develop means of measuring progress in habit and attitude formation along the lines indicated in the concrete acts proposed above, we would soon develop a technic of producing the desired results in boys and girls with a great deal more certainty, ease, and rapidity than we now have reason to believe we shall develop it in the next fifty years. At least the lists above furnish the basis for further experimental work.

The facts in regard to the actual ratings given teachers from 1916 to 1919 are shown in Tables III, IV, and V.

If the marks given below represent the facts, the first semester the rating scheme was in use the median teacher in Republic was a low-average teacher. The second semester, an average teacher; the third semester, an average teacher; the fourth semester, a high-average teacher; the fifth semester, an above-average teacher; and the sixth semester, an above-average teacher. Reference to the rating scheme above would indicate that the gain was from a median teacher of ability 35 to a median teacher of ability 80. In other words, in the judgment of the superintendent, the efficiency of the average teacher in Republic schools as a result-getter *more than doubled* in the three years under discussion, and this in spite of the fact that the above-average teachers were being drawn off by schools which paid better salaries at least as fast as the superintendent was able to improve or eliminate the teacher of below-average ability. The explanation can be found only in the fact that, with better salaries, and clearer notions of what good teaching really meant, the superintendent was enabled to recruit from other cities and from normal schools and colleges a higher grade of teacher material than ever before. For example, 20 percent of the teachers now employed in Republic achieved honors for scholarship in normal school or college. Some others were recommended as the best students in special departments. However, this does not discredit the value of the rating scheme as a supervisory device for the superintendent proved able to develop teachers while in service. Thirty-seven percent of the teachers who have been employed for a semester or more in

TABLE III. RATING OF INDIVIDUAL TEACHERS—
SIX SEMESTERS, 1916-1919

REPUBLIC TOWNSHIP SCHOOLS, REPUBLIC, MICHIGAN

| Teacher | 1916-17 | | 1917-18 | | 1918-19 | | Reason for Withdrawal |
|---------|--------------|--------------|--------------|--------------|--------------|--------------|-----------------------|
| | 1st Semester | 2nd Semester | 1st Semester | 2nd Semester | 1st Semester | 2nd Semester | |
| 1 | C | D | D | D | C | B | |
| 2 | C | C | C | C | C | C | Request of supt. |
| 3 | B | B | B | B | A | A | |
| 4 | C | C | C | B | | | Promotion |
| 5 | D | D | C | C | B | B | |
| 6 | C | C | C | C | C | C | Request of supt. |
| 7 | C | C | D | | | | Request of supt. |
| 8 | B | B | | | | | Poor health |
| 9 | E | D | C | C | C | B | Request of supt. |
| 10 | C | B | | | | | Promotion |
| 11 | D | D | E | | | | Request of supt. |
| 12 | D | C | | | | | Dissatisfaction |
| 13 | D | D | | | | | Dissatisfaction |
| 14 | B | B | B | B | B | B | Retirement |
| 15 | E | | | | | | Marriage |
| 16 | B | B | B | B | A | A | |
| 17 | C | C | | | | | Marriage |
| 18 | C | | | | | | Promotion |
| 19 | C | C | C | C | C | C | |
| 20 | C | C | | | | | Promotion |
| 21 | C | | | | | | Death in family |
| 22 | | C | D | E | | | Request of supt. |
| 23 | | B | | | | | Promotion |
| 24 | | C | B | B | | | Promotion |
| 25 | | | B | B | B | B | |
| 26 | | | E | D | B | B | |
| 27 | | | C | B | B | B | |
| 28 | | | B | B | B | | Poor health |
| 29 | | | C | B | C | B | |
| 30 | | | C | B | B | B | |
| 31 | | | C | C | | | Promotion |
| 32 | | | | C | | | Promotion |
| 33 | | | | D | B | B | |
| 34 | | | | | | B | |
| 35 | | | | | D | D | |
| 36 | | | | | B | B | |
| 37 | | | | | B | B | |
| 38 | | | | | B | B | Death |

A blank in the column "reason for withdrawal" means that the teacher is still in service in Republic schools.

"Promotion" means taken by some other city at a higher salary.

TABLE IV. DISTRIBUTION OF MARKS GIVEN TO TEACHERS,
BY SEMESTERS, 1916-1919
REPUBLIC TOWNSHIP SCHOOLS, REPUBLIC, MICHIGAN

| Mark | 1916-17 | | 1917-18 | | 1918-19 | | Total |
|-----------|--------------|--------------|--------------|--------------|--------------|--------------|-------|
| | 1st Semester | 2nd Semester | 1st Semester | 2nd Semester | 1st Semester | 2nd Semester | |
| A | | | | | 2 | 2 | 4 |
| B | 4 | 6 | 6 | 10 | 11 | 14 | 51 |
| C | 11 | 10 | 10 | 7 | 7 | 3 | 48 |
| D | 4 | 5 | 3 | 3 | 1 | 1 | 17 |
| E | 2 | | 2 | 1 | | | 5 |
| Total.... | 21 | 21 | 21 | 21 | 21 | 20 | 125 |

TABLE V. DISTRIBUTION OF LAST MARKS GIVEN TO TEACHERS
WITHDRAWING FROM REPUBLIC SCHOOLS

| MARK | REASON FOR WITHDRAWAL | | | TOTALS |
|-----------|-----------------------|-----------|---------------|--------|
| | Required | Promotion | Other Reasons | |
| A | | | | |
| B | 1 | 5 | 3 | 9 |
| C | 3 | 4 | 2 | 9 |
| D | 2 | | | 2 |
| E | 2 | | 1 | 3 |
| Total.... | 8 | 9 | 6 | 23 |

Republic schools during the last three years have shown enough improvement to cross the boundary between a lower and a higher rank. Fifty-three percent of the teachers now in Republic schools have improved at least one division. In other words, supervision, as it found expression in suggestions for improvement in the items of the rating scheme, was about equally responsible with elimination of the less capable and the employment in their places of the more capable type of teacher, for whatever improvement has taken place in the efficiency of the teaching corps.

Of course, the question of the accuracy of the ratings of the teachers is an important one. The study has the limits imposed by the fact that one judge, and only one, rated the work of one group of teachers, and only one. On the other hand, the methods

used in rating were, it is believed, vastly more objective and hence, more reliable, than those presented in any other rating study yet made. To the mind of the author the greatest evidence against the validity of the absolute rating is the fact that the median teacher is rated *above average*, and that the distribution of teachers is around that point instead of around a teacher rated *average*. Every other bit of evidence points to the fact that a very marked improvement has taken place in the teaching corps in Republic schools. Hence, if serious error has been made, it was made when the teachers of Republic were rated for the first time in January, 1917.

The absolute authority placed in the hands of the superintendent to eliminate teachers had so much to do with whatever success attended the use of the payment-for-merit plan, that it would be foolish to recommend the plan for general adoption where the authority mentioned does not exist. The reasons for which teachers could be dismissed were as follows:

1. Low efficiency, without, in the judgment of the superintendent, prospect of improvement in the immediate future.
2. Acts injurious to the interests of the school or the community, repeated after warning by the superintendent.

Republic, like every other community in the land, has its school politicians. These men and women study constantly every feature of school administration, supervision, and teaching, not for the sake of providing better school facilities for boys and girls, but for the sake of gaining personal advantage from any phase of school work which can be put in the wrong light before the public. These persons do not have to be reasonable or consistent in order to be troublesome. Soon after the bonuses were made large enough to attract their attention, they began to demand that the amount of the additional compensation should be saved to the public. They contended that a teacher would teach a year for the basal salary just as willingly and just as efficiently as she would for the basal salary plus the bonus. They contended further that the teacher who was not that sort of an idealist was not fit to teach boys and girls anyhow. Even if the scheme could have been defended successfully from attacks from without, it would still have been in the way when it came to employing new teachers; for the teachers desired could hardly have been employed even for

a sum equal to the basal salary and the median bonus (\$100 a year) combined. It was clearly impossible to secure funds to care for the bonuses in addition to the necessary basal salaries. Hence, the additional compensation-for-merit scheme was abandoned in 1919 and the superintendent and most of the teachers breathed a sigh of relief.

The rating scheme proposed above is still in use as a supervisory device, but, of course, without any relation to the salary schedule. Used in this way, the rating card is a valuable supervisory instrument. Additional evidence of this fact was obtained as follows:

Some time following the fifth semi-annual rating the superintendent was requested to present the results of his inquiry before a state meeting of superintendents and school board members. Accordingly, the following notice was sent to teachers in the largest building in Republic Township:

I have been asked to present the results of my investigation leading to a tentative scheme for rating teachers before the State Association of Superintendents and School Board Members. Before accepting the invitation, I should like your candid opinion as to whether your cooperation in forming the scheme and in being rated by it has affected your work and how. Record your conclusions on this paper without consultation and without signing the paper. No effort will be made to hold anyone responsible for adverse judgments. Indeed, no effort will be made to discover who wrote which paper. However, the opinions may be classified as to the amount of professional thought they reveal before they are presented to the Republic teachers or any other audience.

WM. L. CONNOR
Superintendent

Replies were received in sealed envelopes, shuffled thoroughly by a clerical assistant, and typed for the consideration of the superintendent. They ranged from replies as favorable as the following:

My opinion of the rating scheme is as follows:

First, it aids the teacher to measure her pupils, and find their weaknesses, thereby giving her a chance to help them overcome their trouble, whereby the teacher gets the result required.

Second, it helps the teacher to measure herself and gives her a chance to study the standard by which she is measured.

Third, following the rating scheme as closely as possible helps my pupils to think and helps me to help them.

To those as unsatisfactory as the following:

If my cooperation in forming our rating scheme, and being rated by it, has in any way affected my work in the school I am entirely unconscious of it.

Of the seventeen replies received, twelve were favorable and five adverse or noncommittal. The comments, however, make it clear that a great deal of good came out of the work of rating teachers. On the other hand, it would be difficult to read through the discussions presented without gaining the impression that a great deal of misunderstanding and hard feeling existed. Indeed, enough soreness is in evidence to have disrupted any administration not securely founded. However, the dissatisfaction was directed more against the payment-for-merit plan than against the actual operation of the rating scheme as a supervisory device.

It is safe to say that at least three-fourths of all teachers employed in Republic schools were willing to work under the merit plan, and very seldom did the other fourth openly express dissatisfaction. How much of this seeming content was due to the fact that everybody's salary was going up anyhow (the scheme was accompanied by a general rise in wages in no essential way related to it), how much of it was due to the merits of the rating scheme, how much to the fact that teachers cooperated in forming the plan, and how much to the fact that the superintendent exercised rather unusual powers in employing and discharging teachers, it would be hard to tell.

However, if the superintendent of the Republic schools could secure a living and saving wage for the least experienced and poorest trained teacher in the corps, together with reasonable increases for additional experience and training, up to 75 percent more than the minimum wage, he would recommend that the question of extra compensation for merit be reopened.

In the meantime, the rating card proposed above, while imperfect in certain respects, is still in use in supervision, and really represents an advance over previous efforts. It is a useful supervisory device because it directs attention effectively to important elements in the changes which it is a teacher's task to produce in children.

REPRODUCTION AS A MEASURE OF READING ABILITY¹

ARTHUR W. KALLOM

Department of Educational Investigation and Measurement, Boston, Massachusetts

THE IMPORTANCE OF SILENT READING

Reading in some form occupies a large share of the curriculum of both lower and upper grades. In the first grade the time is almost wholly occupied in teaching children how to read; and, oral reading predominates. As the child progresses through the grades the time devoted to oral reading gradually diminishes and the time devoted to silent reading increases.

These two forms of reading have entirely different aims. Oral reading above the primary grades aims to convey the thought to some hearer. Correct pronunciation, clear expression, and good enunciation are, therefore, necessary in the oral work. These, however, are only the mechanics which enable the reader to convey the thought which he has in mind with force and understanding.

In silent reading, on the other hand, the aim is to teach the pupil to obtain the thought from the printed page for himself. This silent reading is followed by some form of reproduction either at once or at some future time. Until recently the teacher of the upper elementary grades has not felt the importance of *teaching* silent reading. Even now it is very probable that she does not realize the full importance of this teaching. Reading, oral or silent, means the recognition of the printed word as a symbol and a correct interpretation of the symbol into a picture for which the symbol stands.

FACTORS IN SILENT READING

This is not by any means a simple process. It may be analyzed into the following factors:

1. Correct visualization of each word
2. Knowledge of the various meanings of each word
3. Choice of the correct meaning as shown by the context

¹ An address delivered before the National Association of Directors of Educational Research at Cleveland, Ohio, February, 26, 1920. In the absence of Mr. Kallom the address was presented by Miss Harriet M. Barthelmess.

4. Forming of the correct relations between these meanings in order to interpret phrases and clauses

5. Forming of the correct relations between phrases and clauses in order to interpret sentences and paragraphs.

This interpretation of words, phrases, sentences, and paragraphs depends on at least two things: (a) the pupil's experience; (b) his ability to recall the experience and apply it to the case at hand.

Any inability on the part of the pupil to make application of his experiences or the making of a wrong application results in error.

Two illustrations will make this thought clear. Question 14 of Kelly's Kansas Silent Reading Test II reads as follows:

A list of words is given below. One of them is needed to complete the thought in the following sentence: The roads became muddy when the snow

Do not put the missing word in the blank space left in the sentence, but put a cross below the word in the list which is next above the word needed in the sentence.

water
is
melted
snow

This question is one of the most difficult in the entire test, being given the value of 4.9. To solve this question correctly the pupil must:

1. Refrain from filling in the blank with the correct word
2. Correctly interpret the phrase "below the word in the list"
3. Correctly interpret the clause "which is next above the word needed."

These three phases must be correctly interpreted not only individually *but also with relation to each other* in order to get the correct answer. The high value of 4.9 for this question is probably due to the fact that pupils do not see the relation between the parts correctly. Pupils do not, as a rule, fail in reading the words of this selection nor in knowing the meaning of the individual words, for none of them is difficult, but they do fail in *understanding* the relation between the words, phrases, and clauses. The failure to see this relation is often laid at the door of carelessness. This may be a cause at times, but more often a lack of experience in meeting such a complicated situation is a contribu-

tory cause and at times may be the important cause of the error. The pupil has been in the habit of placing a cross before or after the word, and when he reads "put a cross" his experience leads him to put a cross before or after the word, thus making an error.

But pupils also make wrong interpretations of individual words. One of the tests given to grade v in May 1919 was the part of the story of Aladdin in which the magician directs Aladdin to go to the bottom of the stairs into a great underground building divided into three vast halls. In each of these halls Aladdin was to find four large brass vessels. A boy who had read this selection interpreted the word "vessels" to mean ships. He had a perfect picture of this underground room as some sort of museum in which these ships were placed. After some questioning as to what other meaning the word "vessels" might have, he remembered the meaning of an "urn" or "vase." Immediately his whole picture was changed. A pupil left to himself, however, does not get a chance to have this impression changed. He may, therefore, go on for some time on the basis of the wrong interpretation.

The teacher may learn of this wrong interpretation only through some form of reproduction. The great criticism against reproduction has been that it involves memory. What is generally meant is word memory without thought or without meaning. This should be decried everywhere and at all times. But the right kind of reproduction, depending on the right kind of memory should be cultivated, for every discussion depends on the ability to think correctly.

BOSTON'S PROGRESS

In 1915 and 1916 Mrs. Ellor Carlisle Ripley, assistant superintendent of schools, was placed in charge of the reading throughout the schools of Boston. In undertaking to improve both oral and silent reading four plans were used:

1. Asking principals and teachers to assemble the children in the auditorium, having the best readers act as an example to others by reading selections before the audience of the other pupils
2. Asking students in the various schools of elocution to read selections suitable for children before the entire school in the auditorium

3. Selecting two or three of the best readers from a number of different schools, assembling them at a central school, and having them read prepared and unprepared selections before an audience of other pupils, parents, and friends

4. Visiting a large number of rooms to hear the children (especially the poorer ones) read in their regular classes.

In conjunction with the third phase of this work, viz., when selected children had gathered from a number of schools in one auditorium, the Department of Educational Investigation and Measurement used the Kelly Silent Reading Test. One hundred and eighteen of these selected children were tested in all. They came from all grades, a few even as low as the fourth, although principals were encouraged to send upper rather than lower grade children. The results showed that these pupils had a large range of individual differences. There was a great deal of overlapping but the medians were from two to eight points higher than the standards set by Professor Kelly.

SELECTION OF MATERIAL FOR SUITABLE TESTING

As a result of having the Kelly test given to this small group of selected pupils, many of the teachers felt that reading ability was not being tested. They, therefore, wished a further test to be made with selections more nearly like the ordinary reading of pupils. As a result of this, in the fall of 1916 a committee of teachers representing grades IV to VIII inclusive, met for the purpose of selecting suitable paragraphs for testing grades IV, V, VI, VII, and VIII. The following directions were given for guiding the teachers in making their selections:

1. That the paragraph selected should be from 100 to 150 words in length

2. That the paragraph should as far as possible tell a complete story

3. That the teachers should select paragraphs suitable for their own grades

4. That paragraphs selected should not be familiar to the children.

As a result of the work of this committee, the department received for each grade four or five selections which the teacher of that grade thought were suitable. The members of the depart-

ment undertook to make from this collection of material a selection of two paragraphs for each grade, which would eventually be used for testing. Each of these paragraphs was given to approximately five hundred children. That is, a paragraph selected for grade iv by a fourth-grade teacher was given to grade iv, a paragraph selected for grade v was given to grade v, and so on. In no case was the paragraph selected for a given grade given to any other grade.

GIVING AND SCORING THE TESTS

The tests were given by a group of normal school seniors who had been trained to give them in a uniform manner. The pupil was given a sheet on which was printed the selection to be read. After this selection had been read once, and the time of reading recorded in seconds, it was collected and a four-page folder given to each pupil upon which he was asked to write a reproduction of the paragraph. After this reproduction was written, he was asked to turn to the last page of the folder and answer some questions based on the paragraph which had just been read. In other words, the pupil was required to do four things:

1. To read the selection once
2. To record the time
3. To reproduce on another sheet all that he could remember
4. To write from memory answers to questions based on the paragraph.

The scoring of the reproduction was based on the number of ideas reproduced. What constituted each particular idea was decided upon previous to the correction of the test. In correcting the material it was very difficult at times to determine whether the idea had been reproduced by the pupil or not. In many other instances it was perfectly clear that the pupil, instead of reproducing the paragraph which he had just read, brought into the reproduction many ideas based upon his experience.

In correcting the reproduction wrong spelling was not considered as an error. How much the inability to spell a word interferes with the ability to reproduce is not known.

RETURNS FROM GRADE IV

The following was one of the stories given the fourth grade to read and reproduce:

Early one morning Washington left his quarters and went alone to see for himself what his soldiers were doing in a camp which he had ordered to be fortified. The weather was so cold that he wore a long overcoat with a great cape. The coat covered his uniform, and his hat and cape did not leave much of his face to be seen. For this reason the soldiers who saw him did not know that the tall man passing by was their general, George Washington.

The following reproductions of the foregoing story illustrate how pupils weave into their reproduction extraneous ideas which have come to them through their experience.

When Georges Washington was a boy he told a lie when grew up to a Gengral no one didn't no him he had a uniform on him and a cape and a cap.

Early one morning Goeoge Washtongton went out and saw Abraham Lincoln in the window and said to him do you no me yes I do no you George Washtongton.

In the first instance the pupil retains the ideas of "George Washington," "a general," "uniform," "cape," and "cap." In addition he has added the ideas of "when he was a boy he told a lie." In the second reproduction the only thing which was retained was that the story was about George Washington. Bringing in an idea, such as that of Lincoln, which has nothing to do with the situation is an interesting but not uncommon occurrence.

Attention has been called to the fact that one source of error is a misinterpretation of the meaning of individual words. The two following reproductions of the same story show the misinterpretation of the word "quarters":

One morning Washington left his *quarters* on the table and went to see how his solderes were dowing.

Once Goerge Washington said he would go out to see what his *quarters* where doing in the camp. It was cold so he had to put on his big warm coat. He saw what his doildiers were doing. He said let us go out to camp.

The first pupil shows that of the different meanings of the word "quarters" he has selected the wrong one. He may have done this either through ignorance or through making a wrong application of his experience due to failure to remember the relation between the phrases in the selection given. The second pupil evidently has a very hazy idea of the meaning of this word and possibly no correct meaning for it.

RETURNS FROM GRADE VI

Below is given one of the selections given in grade vi under the same conditions and directions as used in grade iv.

The baby seals make us think of young dogs and they play together like puppies. The mother's call to her little ones is similar to that of a sheep, and her babies' cry is like the bleating of a lamb. When they are about six weeks old they go into shallow pools and learn to swim. As they grow stronger they roll down into the sea and take excursions of miles. The seals are so valuable that men are allowed to kill only a few of them each year. They are not hard to catch. The seal hunters carry clubs about seven feet long and as thick as a baseball bat. They stun the animals by striking them on the skull and then other men come along with sharp knives and kill them.

The following is a reproduction written by an eleven-year-old girl:

The baby seals resemble our dogs and puppies. Their mother's cry sounds like a sheep. When they are old they are put in a tank to learn to swim. When they are older they toss balls in a circus.

This reproduction indicates the importance of previous experience. Such a piece of work is often discarded by the teacher as useless. As a matter of fact it is a valuable means of determining what should be done for the pupil. The question to be answered is: "Why does the pupil bring in the ideas about the tank and the circus?" It is the answer to this "Why" that should guide the teacher in her future corrective work with the pupil.

QUESTIONS

As noted above, a second part of the test consisted of a number of questions based on the selection read. The first question for the fourth grade, with the answers which the department deemed acceptable and those which it deemed non-acceptable, is given below. Each of the non-acceptable answers was given by one or more pupils. In all, 42 percent answered the question correctly.

"Why did Washington leave his quarters?"

Acceptable

to see his soldiers
to see soldiers in
camp
to see his men
to see his army.

Not Acceptable

because of his trail, to see a man, to go on
an errand, on important business, it was
cold, he was a general, he did not want to
take it with him, he didn't want to be in
there, he had too much, they would not

know him, he died, he did not like it, his mother died, to look at himself, to see what they were doing, because he had to, the quarters are American money, he left it in his pocket, because he wanted to, because they told lies, to go to fight the Germans, to go to a 'feet'.

Three types of answers are to be noted among the non-acceptable:

1. Those which apparently come from another part of the story

Illustration: because it was cold

because he was a general

2. Those which apparently result from a failure to use the right meaning of a word

Illustration: the quarters are American money

he left it in his pocket

3. Those which apparently have nothing to do with the story but are based on outside experience

Illustration: because his mother died

because they told lies

to go to fight the Germans.

From the returns as a whole it was plain that the paragraph was too difficult for fourth-grade pupils to read and get the thought. Only 5 percent answered correctly the question "What had Washington ordered done?" Neither was this idea included generally by the pupils in their reproduction.

In answering the questions on the story about the seals the sixth-grade pupils showed the same tendency to draw on their experience as was exhibited by the fourth-grade pupils. At one of the parks in Boston there is an aquarium where there are a number of seals. The familiarity of children of the city with this aquarium was shown by the pupils of the sixth grade in their reproductions, and particularly in their answers to the fourth question: "Where do they learn to swim?" when they wrote "at Marine Park."

Below are given the different answers to the first two questions of this selection—answers which seem to indicate again that the pupil draws largely on his experience to answer questions. If the experience suggested by the reading is more vivid than the mind

picture he gets from correctly interpreting the selection read, he tends to tell the story of that experience.

"What do baby seals look like?"

Acceptable

Not Acceptable

dogs
puppies.

snails, beavers, bears, long and narrow, head of a rat and tail of a fish, fox, face like a rat and tail like a fish, gold, like tadpoles, baseball bat, crabs, like big ones, bat, smooth soft skin, dark fur, cat, fish, very small, mice, monkey, lamb, long fur with spots for eyes, little cubs, Eskimos, tiger, hunter, dragon, walrus without tusks.

"What does the young seals' cry sound like?"

lamb
bleating o sheep.

dog, cat, baby, loud cry, bird's cry, cow, lion, horn, bear, screech, deer, whistle, horse, dog when he watches, very big, like a ten-cent horn, saw, winds, like a wake, groan, elephant, ma ma, goat, sweet, squeaky, hoarse voice, harsh, very loud noise, echo, walrus, shriek, wolf, girl, when you sleep, whales, winds.

The first question was answered correctly by 70 percent of the group and the second by only 56 percent. Many answers show that pupils did not even approximate the correct answer and yet the answer shows an intimate knowledge of the seal as an animal at the aquarium. One who is familiar with the seal would not find it difficult to make most of the descriptions fit.

SUMMARY

One of the purposes of a standard test is to give the teacher information concerning her class and its individual needs. Although it is necessary to have reproduction, oral or written, in most branches of the curriculum of the upper grades as a means of judging whether the pupil has learned the lesson assigned, as a basis for standard tests it will probably never be practical. Such a test violates two important rules of educational measurement in that (a) the test is subjective rather than objective and (b) the task of correcting results is prodigious.

On the other hand, answers to carefully selected questions yield as good a measure of pupils' ability to read as does reproduction, and the answers are easier to score.

In spite of these objections, however, there are certain advantages in giving such a test, for it reveals that:

1. Teachers cannot be depended upon to decide accurately whether a selection is difficult or not—This should make teachers more lenient in their criticisms of the results of the pupils' study

2. It is the relation between words, phrases, and sentences which makes the difficulty rather than the meanings of the individual words, phrases, and sentences

3. There will be certain ideas in the selection that very few pupils will be able to reproduce—These may be important or not

4. The pupil will draw very largely upon his experience and not be bound down by what he has just read. This gives the teacher a clue as to how corrective material should be used.

THE USE OF PSYCHOLOGICAL TESTS IN THE EDUCATIONAL GUIDANCE OF HIGH-SCHOOL PUPILS

W. M. PROCTOR

Leland Stanford Junior University

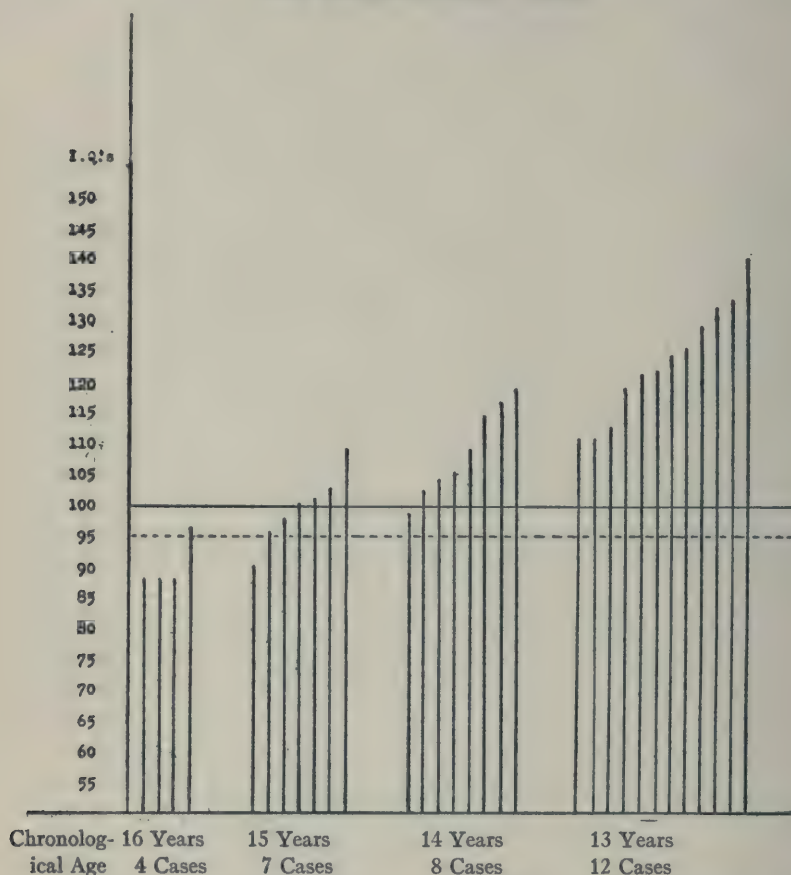
In a previous article¹ the writer has shown that there is a very close relationship between intelligence level and the elimination of pupils from high school. There are good grounds for the prediction that 75 percent of those who test below average, mentally, will fail in more than one-half of their studies during their first year of high school; that 50 percent of them will leave school to go to work during the first two years; and that none of them will remain to graduate. The members of this group demand special attention. If they are not carefully guided in the selection of their high-school work many of them will fail in all of their subjects during the first semester, with elimination as the inevitable result.

Heretofore we have acted upon the assumption that the only way to discover the capabilities of high-school pupils was to permit them to try themselves out in various subjects of the high-school curriculum. It has even been a common practice to require all first-year high-school pupils to take an abstract subject like algebra. Thus many a promising educational bark has gone down on the uncharted rocks of the first-year high-school subjects.

Experimentation with psychological tests has now reached a stage where we can begin to hope that this wasteful trial and error method of procedure is to be discarded for a more scientific plan of educational guidance. No tests have yet been devised, either mental or pedagogical, which enable the adviser to chart unerringly the educational possibilities of a given youth, but we have made sufficient progress in the science of mental measurements to be able to estimate with approximate accuracy the probable school success of a given pupil. The results of a practical experiment in educational guidance by means of mental tests will be described in the following pages.

¹ Proctor, W. M., "Psychological tests as a means of measuring the probable school success of high-school pupils," *Journal of Educational Research*, 1:258-70, April, 1920.

FIGURE 1.—ILLUSTRATING DISTRIBUTION OF 8A² PUPILS BY I. Q.'S AND CHRONOLOGICAL AGES



1. *Conditions surrounding the experiment.*—The mid-year 8A class of the Palo Alto Intermediate School comprising 31 pupils, was examined with the Stanford-Binet Scale in January, 1918. Figure 1 shows the distribution of intelligence quotients by chronological age groups. There were twelve thirteen-year-old pupils, eight fourteen-year-old, seven fifteen-year-old, and four sixteen-year-old pupils examined. The lowest I. Q. discovered in the thirteen-year-old group was 110. All of the thirteen-year-olds

² In grade designations "A" means "second half."

were, therefore, accelerated mentally. None of the fourteen-year-old group fell below 95 I. Q. They were all at or above age mentally. One fifteen-year-old had an I. Q. of 90 and would be listed as "below average" in intelligence. Of the four sixteen-year-old pupils, all of whom would be accounted chronologically retarded, three had I. Q.'s of 88, and one an I. Q. of 96. That is to say, one of the sixteen-year-olds was normal and three would rate as "dull-normal."

Of the 31 members of the 8A class, 22 entered the Palo Alto High School in February, 1918. Four members of the class entered secondary schools elsewhere, two entered business colleges, two went to work, and one, a girl with an I. Q. of 88, eloped with a soldier from a nearby encampment.

Only four members of this 8A class had earned I. Q.'s below 95. One of these, as just indicated, did not enter high school. The remaining three entered high school, but two of them dropped out before the end of the first year. The percentage of elimination from the class of those who tested below 95 I. Q. was therefore 75.0. The one who remains in high school is making an average record in scholarship. She made a rating of "average" on an Army Scale test given a few weeks later, and has shown herself capable of diligent application to her school work.

Each member of the class was questioned as to his or her vocational ambition, educational plans, and the subjects which he or she would like to take up during the first year of high school. These data, as well as results attained in Stanford-Binet and Army Scale mental tests, were recorded on cards.

2. *Use made of information gathered.*—The cooperation of the vice-principal and entering class adviser was obtained in order that the members of the 8A class might receive intelligent educational guidance when they entered the high school. Duplicate sets of cards were prepared for the use of these persons. On registration day all the newly entered 8A graduates were referred to the vice-principal or the class adviser for help in making out their curriculums. No one was permitted to file a study card without this conference.

Following are samples of the cards used in this experiment in educational guidance. The only change from the original is in the case of the name.

CARD NO. 1

| | |
|--|-----------------------------|
| Smith, Jane | Chronological age: |
| | 13 yrs., 9 mo. |
| Score Army Scale...151 | Stanford-Binet mental age: |
| Army Scale mental age: | 16 yrs., 3 mo. |
| 17 yrs., 5 mo. | |
| Army Scale I. Q.....124 | Stanford-Binet I. Q.....118 |
| | |
| High school subjects which pupil desires to take: | Educational plans: |
| Sewing | To finish high |
| French | school and take |
| History | business college |
| Typewriting | course. |
| Piano | Vocational ambition: To be |
| | a stenographer or |
| | bookkeeper. |
| | |
| Grade of work done in intermediate and grammar schools: | |
| "B" and "B+" | |
| | |
| Comment of Examiner: If assigned to algebra can safely be placed in first "rapid progress" division. | |

Jane Smith was advised to take subjects that would make it possible for her to go on to college as well as to carry out her ambition to become a stenographer. She took up: English German, algebra, and typewriting. During the first semester of her high-school course she made an average of "B" in all of her subjects.

Card No. 2 relates to Mary Jones who took exactly the subjects outlined on her card, except that being a first-year pupil she was permitted to take four subjects only and had to wait until her second year for drawing. At the end of the first semester in high school she had earned three "A" grades and a "B+," with an average of 93 percent.

Here was the case of a girl with very superior ability as indicated by two different mental examinations, by her school record and by the estimates of her elementary and intermediate school teachers. She gave evidence of being an independent thinker, of knowing

CARD NO. 2

| | |
|--|--|
| Jones, Mary | Chronological age: 13 yrs., 1 mo. |
| Score Army Scale...148 | Stanford-Binet mental age: 17 yrs., 2 mo. |
| Army Scale, mental age: 17 yrs., 2 mo. | |
| Army Scale I. Q....131 | Stanford-Binet I. Q.....131 |
| High school subjects which pupil desires to take: | Educational plans: To finish high school and attend a university or normal school. |
| English | |
| Algebra | |
| Latin | |
| History | Vocational ambition: To become a drawing teacher. |
| Drawing | |
| Grade of work done in grammar and intermediate schools: "A" | |
| Comment of Examiner: Knows just where she is going and how to get there. May safely be permitted to select her own course of study. Assign to first division in algebra. | |

just what she wanted to make of herself, and just what she would have to do by way of preparation. The necessity for educational and vocational guidance in her case might well be questioned. However, it was a real advantage to her teachers to know at the very beginning of her high-school career the quality of her ability and something of her life plans in order that they might give immediate and sympathetic cooperation. Without this knowledge they might have made the mistake of holding her back to the pace of the "average" pupil. Fortified by the facts relative to her mental gifts and vocational ambitions, she is to be permitted to complete her high-school course in three years.

There is just as much danger that the bright pupil will not be given enough to do, as that the dull pupil will be given tasks that are too difficult to perform.

CARD NO. 3

| | |
|--|---------------------------------------|
| Roe, Richard | Chronological age: 14 yrs., 4 mos. |
| Score Army Scale...150 | Stanford-Binet mental age: |
| Army Scale mental age: 17 yrs., 4 mo. | 16 yrs. 9 mo. |
| Army Scale I. Q....120 | Stanford-Binet I. Q.....117 |
| | |
| High school subjects which pupil desires to take: | Educational plans: |
| English | To finish high |
| History | school then at- |
| Algebra | tend a university |
| French | or the U. S. |
| | naval academy. |
| | Vocational ambition: Chemi- |
| | cal engineer or |
| | naval officer. |
| | |
| Grade of work done in elementary and intermediate schools: | |
| Very poor. Estimated as "average" | |
| by some grade teachers, and as | |
| "below average" by others. | |
| | |
| Comment of Examiner: Boy has ability but | |
| needs to be waked up. Suggest | |
| that he take general science in | |
| place of history for first | |
| year. Also suggest that he be | |
| placed in first division in alge- | |
| bra where he will have to work. | |
| He will need to develop ability | |
| in both science and mathematics if | |
| he is to follow his vocational | |
| ambition. | |

The boy whose card is set forth above enrolled for English, algebra, history, and general science, upon entering high school. During his first half year he made grades of "C" in English, and "B" in each of his other subjects. This was a great improvement over the grades earned by him in the eighth grade. A recent

checking up shows that he has brought up his English grade and is maintaining college recommendation standing in all his work. The mental tests were an aid to his advisers in discovering how to spur him on to creditable achievement in his school work.

CARD NO. 4

| | |
|--------------------------|------------------------------|
| Brown, Carrie | Chronological age: 15 yrs., |
| Score Army Scale... 100 | 7 mo. |
| Army Scale mental age: | Stanford-Binet mental age: |
| 14 yrs., 0 mo., | 14 yrs., 2 mo. |
| Army Scale I. Q. 89 | Stanford-Binet I. Q. 90 |

.....
High school subjects which
pupil desires to take:

English

Algebra

Latin

Typing

Drawing

Educational plans:

To go to Mills

College

Vocational ambition: To be
a Chemist.

.....
Grade of work done in intermediate and grammar schools:
Grades in 8A class only fair, even
in work that is being repeated.
Estimates of elementary and inter-
mediate teachers: "slow" but
a conscientious worker.

.....
Comment of Examiner: Should be discour-
aged as to taking Latin. Algebra
doubtful, but if she insists in
view of desire to go to college,
assign to second division.

The program finally worked out by Carrie Brown and the class adviser included English, algebra, free-hand drawing, and typing. Her grades at the end of the first semester in high school were: English, "C"; algebra, "E" (failure); free-hand drawing, "C"; typing, "B." She had failed in algebra, the subject counted as doubtful by the examiner, and had earned less than college recommendation grades in the only other subjects that would be counted

toward entrance to Mills College. Even if she completes high school her chance of continuing in college is not at all bright.

A sufficient number of sample cards have been described to illustrate the method employed. There was no coercion. Counsel and advice in the selection of subjects were all that was attempted, but the counsel and advice offered were based on all the significant information with regard to mental ability, school success, vocational ambition, and teachers' estimates of ability, that could be obtained. Certain pupils elected to take subjects which the advisers felt sure they would fail in and made passing grades. Others taking subjects on the advice of the counselors failed. Such cases simply illustrate the truth that no human agency, however fortified with information or however careful to mix common sense with theories, can hope to be infallible. The general results of the experiment, however, were very satisfactory.

3. *Results attained by "guided" and "unguided" high-school pupils.*—The original group of 107 high-school pupils which entered the Palo Alto High School in September 1916, were examined with the Stanford-Binet Scale after entering high school. They had already selected their courses of study at the time of being examined, hence can properly be designated as "the unguided" group. A comparison of the first year's work done by the members of the "unguided" group with the work done by the group that entered high school in February, 1918, and which had the benefit of counsel based on mental tests and other significant data, will illustrate the value of careful guidance as against the trial and error method of selecting high-school courses of study.

TABLE I. COMPARATIVE FACTS REGARDING "GUIDED"
AND "UNGUIDED" GROUPS OF HIGH-SCHOOL PUPILS

| Group | Out at Work | Per- cent | Out by Transfer | Per- cent | Failed 1 Subject | Per- cent | Failed 2 or More | Per- cent |
|------------|----------------|--------------|--------------------|--------------|---------------------|--------------|---------------------|--------------|
| Guided.... | 1 | 4.5 | 2 | 9.0 | 4 | 18.0 | 0 | 0.0 |
| Unguided.. | 13 | 12.0 | 14 | 13.0 | 33 | 31.0 | 11 | 10.0 |

It is not exactly true to fact to designate the original group of 107 as the "unguided" group. Most of them had been examined with the Stanford-Binet before the end of the first six-week period. Whenever the results of the first six weeks of school work con-

firmed the indications of the mental tests that a pupil would probably fail in such abstract subjects as algebra, Latin, etc., that pupil was permitted to drop the subject and continue the semester carrying but three subjects. The subject dropped at the end of the first six-week period was not counted as a failure in compiling the data for Table I. The mental tests were utilized to aid in correcting the mistakes made by the pupils in the unguided selection of their subjects. To this extent, then, the original group was guided, but the guidance came after, not before school work was begun.

The number of failures registered against the "unguided" group at the end of their first year in high school would undoubtedly have been greatly increased if it had not been for the limited guidance above described. This fact gives greater significance to the data presented in Table I.

The median I. Q. of the unguided group was 105, and the median I. Q. of the guided group was 108, a difference of three points in favor of the guided group, but this is not in itself an adequate explanation of the superior record made by the guided group during its first year in high school.

The most significant facts to be found in Table I are those relating to number of failures in one subject, and in two or more subjects. It appears that 31.0 percent of the unguided group failed in one subject, and 11.0 percent failed in two or more subjects during their first year in high school, while of the guided group only 18.0 percent failed in one subject and none of them failed in two subjects. Since it has been shown that failures in school work tend to increase the percentage of elimination it is reasonable to attribute the small percent of elimination due to leaving high school to go to work, in the case of the guided group, to entire absence of failures in two or more subjects.

The results of the above described experiment in educational guidance by means of mental tests would seem to justify the conclusion that such tests may be of material assistance to the high-school administrator, if used in connection with other significant data. It is certain that the methods applied in this instance, if employed in any high school, would prove greatly superior to the wasteful "trial and error" methods that now prevail.

4. *Relation of general level of intelligence to success in a given subject.*—In Table II is shown the correlation between general

level of intelligence and high-school grades in algebra. The algebra grades are recorded under the letters A, B+, B, C, C-, D, and E. The I. Q.'s come under the groups 84-89, 90-94, 95-99, etc.

TABLE II. CORRELATION BETWEEN I. Q.'S AND GRADES
IN ALGEBRA OF 113 HIGH-SCHOOL PUPILS

| Algebra Grades | I. Q.'s (STANFORD-BINET) | | | | | | | | | | Totals |
|-------------------|--------------------------|-----------|-----------|-----------|-------------|-------------|-------------|-------------|-------------|--------------|--------|
| | 80- 84 | 85- 89 | 90- 94 | 95- 99 | 100- 104 | 105- 109 | 110- 114 | 115- 119 | 120- 124 | 125- over | |
| "A"..... | | | | 2 | 1 | 3 | 0 | 1 | 1 | 3 | 11 |
| "B+"..... | | | | | | | 1 | 2 | 2 | 5 | 10 |
| "B"..... | | 2 | 2 | 3 | 7 | 3 | 7 | 5 | 5 | 2 | 36 |
| "C"..... | | 1 | 4 | 5 | 6 | 2 | 5 | 3 | 3 | 3 | 32 |
| "C-"..... | | | | 1 | | 1 | | | | 1 | 3 |
| "D"..... | | 1 | | 2 | 1 | 1 | 1 | | | | 6 |
| "E"..... | 1 | 3 | 5 | 3 | | 1 | 2 | | | | 15 |
| Totals..... | 1 | 7 | 11 | 16 | 15 | 11 | 16 | 11 | 11 | 14 | 113 |

Correlation (Pearson): 0.46

Probable error: 0.05

The correlation obtained, 0.46, which is nine times the probable error can be counted as having considerable significance. An examination of the data contained in Table II will show that twelve of the fifteen failures in algebra, or 80.0 percent of the total number of failures, were earned by pupils with I. Q.'s below 100. There were 35 pupils having I. Q.'s below 100, and 26 of them, or 74.3 percent earned marks below "B," which is the college recommendation grade in California. On the other hand there were 78 pupils with I. Q.'s above 100, and 45 of these, or 63.0 percent, earned marks of "B" or above. An I. Q. of 120 or over denotes very superior intelligence. There are 25 cases with I. Q.'s 120 or over in Table II. All of these succeeded in earning passing grades in algebra. Only one received a mark as low as "C-."

High-school statistics show that algebra is responsible for more failures of first-year pupils than any other subject. Table II suggests that probable success or failure in algebra can be inferred from the general level of intelligence as disclosed by mental tests. It follows that the high percentage of failure in algebra could be

materially reduced if only those were encouraged to take the subject whose general level of intelligence measured up to average or better.

5. *Relation of score in a particular mental test to success in a given high-school subject.*—Test No. 9, in Army Scale, Group Examinations *a* and *b*, is a "word relations" test. It involves a knowledge of word meanings and also the ability to use words intelligently in their proper relations to one another. Table III compares the scores made by 171 first-year pupils of the Palo Alto High School in Test No. 9, with the grades made by the same pupils in English during their first high-school year.

The directions for giving the test are as follows:

In each of the lines below, the first two words have a certain relation. Notice that relation and draw a line under the one word in the parenthesis which has that particular relation to the third word. Begin with No. 1 and mark as many sets as you can before time is called.

Sample sets are then given by the examiner to illustrate what is wanted:

sky—blue: grass—(grow, green, cut, dead)
 fish—swims: man—(boy, woman, walks, girl)
 day—night: white—(red, black, clear, pure)

TABLE III. CORRELATION BETWEEN GRADES IN ENGLISH AND SCORES IN GROUP INTELLIGENCE TEST NO. 9 OF ARMY SCALE, OF 171 HIGH-SCHOOL PUPILS

| Marks in First Year English | Scores in Test No. 9, Examinations <i>a</i> and <i>b</i> | | | | | | | | | Totals |
|-----------------------------------|--|---------|-----------|-----------|-----------|-----------|-----------|-----------|------|--------|
| | 1- 4 | 5- 9 | 10- 14 | 15- 19 | 20- 24 | 25- 29 | 30- 34 | 35- 39 | 40 | |
| "A"..... | | | | 1 | 4 | 1 | 3 | 2 | | 11 |
| "B+"..... | | | 1 | 1 | 7 | 14 | 10 | 2 | | 35 |
| "B"..... | | | 1 | 5 | 16 | 19 | 13 | 6 | | 60 |
| "C"..... | | 3 | 9 | 17 | 17 | 8 | 6 | 2 | | 62 |
| "D"..... | | | | | 1 | 1 | | | | 2 |
| "E"..... | | | | | 1 | | | | | 1 |
| Totals..... | | 3 | 11 | 24 | 46 | 43 | 32 | 12 | | 171 |

Correlation (Pearson): 0.48

Probable error: 0.04

There were forty sets of words and the time allowed was three minutes. For purposes of tabulation the scores made by the

pupils are grouped: 1-4, 5-9, . . . 35-39, 40-. The English grades are indicated by the letters A, B+, B, C, C-, D and E.

The correlation, 0.478, is twelve times the probable error and indicates that a good score in the "word relations" test is a fairly good index of ability in first-year high-school English. The marks in English were the final year marks, which represent the minimum of "D's" and "E's," because all conditions and failures had been removed that could be made up. The median score in Test No. 9, was 26, hence scores of 30 or over could be considered superior. There were 44 pupils who made scores of 30 or over, and 36 of these or 82.0 percent made marks in first-year English of "B" or above. Of the 89 who made scores of from 20 to 29 inclusive, 68.7 percent made marks in first-year English of "B" or above. While of the 38 who made scores of from one to nineteen inclusive, only 23.7 percent made marks of "B" or above.

Stated in terms of recommendation for college the above analysis of the data of Table III means that 76.3 percent of the high-school pupils whose ability in word relations is represented by a score of less than 20 fail to secure a college recommendation grade, while only 18.0 percent of those who attain a score of 30 or over in the same test fail to secure such a grade.

These results seem to indicate that a series of tests involving the fundamental traits of mind essential to the successful study of English could be devised. But there are so many mental traits involved in the mastery of the subject of English that a series of tests bringing into play all those traits would undoubtedly be found to be a good test of general intelligence as well as a test of specific ability in English.

SUMMARY AND CONCLUSIONS

1. The results of an experiment in educational guidance, in which all the members of an 8A class about to enter high school were given mental tests and advised with reference to their first year high-school work, proved very satisfactory. Compared with an unguided group it was found that while 31 percent of the unguided group failed in one subject, and 11 percent failed in two subjects during their first high-school year, only 18 percent of the guided group failed in one subject and none of them failed in two subjects. The mental tests aided in the discovery of the pupil's

general level of intelligence, made possible the giving of sound educational advice at the time when it would do the most good, and thus tended to reduce the percentage of failure and elimination. The methods employed can be adapted to the needs of any high school.

2. The general level of intelligence is shown to have real significance as a means of predicting success in a particular subject, such as algebra. In Table II it was shown that 100 percent of those having I. Q.'s of 120 or over passed in their algebra, while 40 percent of those with I. Q.'s below 100 either failed or were conditioned in algebra, and only 20 percent of the members of this group earned college recommendation grades. The general level of intelligence could be relied upon as a means of selecting those who would be most likely to succeed in algebra and kindred subjects.

3. Success in a particular test of a series may be a fairly reliable index of success in a high-school subject involving the mental traits supposed to be measured by the test in question. High scores in the "word-relations" test of the Army Scale, Examinations *a* and *b*, correspond generally to high marks in first-year high-school English, while scores below twenty in the same test correspond generally to low marks in the same subject.

4. Mental tests for purposes of prognosis in individual high-school subjects such as algebra, English, etc., could no doubt be devised. But the mastery of any high-school subject involves such a complex of mental traits that any test which proves to be a good test of ability to succeed in one subject is quite apt to be found a good test of general mental ability. It has already been shown that a high-school pupil having a high level of intelligence will probably succeed in all of his subjects, and conversely that a pupil having a low level of intelligence is apt to fail in most of his high-school subjects. The best way, therefore, in which to arrive at an estimate of a given pupil's probable success in a specific high-school subject is to discover the general level of his intelligence. It follows that standardized mental tests may render invaluable service in the educational guidance of high-school pupils. Particularly will this be found to be true if the results of the mental test are interpreted in the light of other significant data, such as school marks made in previous grades, teachers' estimates of ability, and educational and vocational plans.

THE OPERATION OF BUREAUS OF EDUCATIONAL RESEARCH—WITH PARTICULAR REFERENCE TO WISCONSIN

W. W. THEISEN

State Department of Public Instruction, Madison, Wisconsin

A remarkable development in scientific methods of studying educational problems has taken place within the past few years—a development which has made it possible for well-informed persons to know more fully and accurately than ever before the possibilities and limitations of educational improvement.

The evidence of this development may be found in many types of material, e.g., in numerous published surveys, school reports, reports of educational conferences, bulletins of bureaus of research, scores of recent educational monographs, and hundreds of magazine articles. Our methods of treatment of the statistics of attendance, progress and promotion, courses of study, teaching personnel, buildings and equipment, and finance have been greatly improved; while more noteworthy than any other has been the advance in the technic of measuring the achievements of pupils.

Unfortunately, however, the utilization of these scientific methods is still far from common in school practice. Moreover, the knowledge of what has been done and can be done in education is still largely the possession of a few. The problem now is to make these methods and this knowledge widely available and practically effective. To accomplish this the results, methods, and spirit of investigation must be put into operation in every school.

The so-called bureau of research is the most active and effective agency today for making these newer developments a part of our practice. Without the intelligent direction and supervision of such an agency, teachers and administrators will not benefit to the fullest extent from our advance in scientific methods. Scientific supervision of this sort enables the weak schools to gain strength and the strong to grow still stronger; makes the successful activities of one school available for others; and develops teamwork among schools. Incomparably more

effective are the united efforts of such schools than the spasmodic efforts of single schools. It is the function of bureaus of school research to provide supervision of this kind, and to bring about over a large area the application of scientific procedure to problems of administration, supervision, and teaching. School systems that have installed research bureaus have come to stand out in marked contrast with others that have failed to do so.

MEASUREMENT OF INSTRUCTION AS A BUREAU PROBLEM

The problems that should command the attention of bureaus of research are many. One of the most important is the problem of applying scientific procedure in measuring and improving the product of instruction. At present only a comparatively few principals and teachers can be expected to proceed unaided. Those who can, will nevertheless lack the stimulation which results from a knowledge of the success of others in wrestling with the same problems. The need of supervision of educational measurement in this restricted sense arises from another source. Many of our school systems, city and state, are deficient in the amount of helpful supervisory aid which they can give to teachers. Educational measurement furnishes a very fruitful instrument of supervision. With a comparatively small expenditure of money, a large number of schools and teachers can be effectively reached. This fact has undoubtedly had its influence on the rapidity with which bureaus of educational measurement, or research, have developed. Such organizations ought soon to be found in every state and in every city of large or moderate size.

Educational measurement on a large scale involves the cooperation of superintendents, principals, and teachers. It is the task of bureaus of research to build up such cooperation. To secure it requires well-directed efforts on the part of those in charge of the bureaus. In the first place, it should be assumed that the large majority of superintendents, principals, and teachers are progressive at heart, that they earnestly desire to know how they can improve the quality of instruction, and that they will decide upon a wise course of procedure, if the facts are clearly presented. In the second place, it will be necessary to make a clear presentation of the merits and possibilities of educational measurements, and in doing so there will be abundant

opportunity for the employment of tact and good judgment. Moreover, the supervisor or director in charge must maintain first-hand contact with the schools through field work. Without it he will not be in a position to realize the possibilities and limitations of local schools and teachers, nor can he prescribe accurately for their improvement. This is exceedingly important in all the bureau work of a state department or university.

The attitude of those whom the director hopes to reach will depend upon their previous acquaintance in the field of educational investigation and measurement. For convenience we may divide school heads and teachers into three groups: Group I: those partly acquainted with educational measurements and their uses, as a result of what they have read or have heard on various occasions; Group II: those wholly, or almost wholly, unfamiliar with the newer developments in educational investigation; and Group III: those who have had considerable experience in their application. The director's efforts to stimulate educational research will differ in each of these cases.

GROUP I. THOSE PARTLY ACQUAINTED WITH EDUCATIONAL MEASUREMENT

Most superintendents even in states which do not support bureaus of research, are partly acquainted with educational measurements and their uses. They are, therefore, for the most part in a receptive mood. Although they are often eager to begin actual investigation, they, nevertheless, hesitate because of unfamiliarity with the details of procedure and lack of confidence in their own ability. In such cases, the presence of the supervisor or director is all that is needed to set the movement on foot.

In visits to schools of this type, the director must adopt a sympathetic attitude. He must endeavor to find out what the schools have tried or are trying to do. By so doing he will discover wherein his own services will be of most help. Among other things, he should make inquiry as to the superintendent's plan for the professional development of his teaching staff. Again, he should learn what subjects of the curriculum are being attacked with a view to improving instruction. The services of the director will usually be most welcomed, if he exhibits a desire to aid in

carrying out the superintendent's plans. With this point as a beginning he will be in a position to make recommendations for additional lines of investigation.

Where such schools have already planned to make some application of educational measurements, they will welcome the director's aid and guidance in the details, such as administering a test, and scoring and tabulating the results. While we do not believe that much of his time should be spent in this fashion, such assistance is often quite necessary in new fields. As far as possible, however, his time should be given to interpreting and teaching others how to interpret the results attained, assisting teachers in the finer diagnostic analysis of results, making suggestions for remedial work, and planning for further experimentation. By all means he should make it a point to meet teachers and principals individually and in groups. The findings and suggestions can then be made the subject of discussion. It is very important that teachers be made to feel that their use of measurements must not stop with a compilation of results, and it is at this point that the services of the director are most needed. Without constructive guidance much well-intentioned effort is apt to count for naught.

Where the supervisor or director finds it impossible to be present in person, he can assist materially through correspondence. When he is asked for information on such matters as school building standards, grading and promotion standards, available tests in various subjects and suggestions as to their use, he is in a position to help superintendents or principals to proceed wisely and to steer them away from pitfalls. In the case of tests, he should ask in each instance that a copy of the results be sent him for examination and criticism.

GROUP II. THOSE UNACQUAINTED WITH EDUCATIONAL MEASUREMENT

Every director, particularly in a new field, must expect to find some superintendents, principals, and teachers who are to all intents unacquainted with educational measurements and their significance. He should endeavor to reach as many of these as possible, thus recruiting the ranks of those better acquainted with the field. There are several avenues and instruments of approach that can be used to good advantage, e.g., conventions of superin-

tendents and principals, teachers' institutes and other educational gatherings and the columns of state educational publications.

The most fruitful reward of effort in this direction results, as before, from personal visitation. This offers opportunity to "talk things over," to tell what has been done elsewhere, to show samples of standard tests, to tell how they may be used, to point out their merits and demerits, to indicate the benefits to be gained from their application, and to suggest lines of investigation. If this part of the work is well done, it will usually result in an invitation to undertake some experimentation, to supply some standardized tests, or possibly to meet with the teachers and explain to them the uses of such tests. New bureaus will find this type of field work advisable whenever it is at all possible.

GROUP III. THOSE EXPERIENCED IN EDUCATIONAL MEASUREMENT

In schools that have already made more or less use of scientific methods of attacking educational problems, the director's efforts can be centered upon an analysis of some of the results and their implications. It should be his business to see, for example, that the diagnostic phase of testing is properly understood and emphasized. It will be a part of his task to demonstrate from the actual papers of the children the diagnostic values and possibilities of the tests used, to suggest forms for tabulating and effectively representing the findings, to indicate possible needed revisions in teaching method, course of study, or gradation of pupils, to propose additional equipment in the way of books, maps, scales for testing, practice materials, etc., and to plan lines of further research and experimentation. Clearly, the director should be a good general supervisor of instruction. He will often find it impossible to give each school all of the personal attention it needs. If, however—as in a state department—there are general supervisors of instruction at work out in the field, they can be instructed and enlisted to aid in carrying out remedial measures in local schools. Their cooperation will be extremely valuable.

Not too much should be expected of schools at the start. In the first place, it is not advisable to attempt many lines of investigation in a given school system, particularly if several of them affect the same teachers. When standardized tests are applied, it is well to confine the testing to one subject at a time. Rarely

should more than two be attempted simultaneously. If too many tests are undertaken, confusion results, and the satisfactory treatment of any one of them is difficult. The professional energies of the teaching staff should be organized for a "drive" on the subject chosen. Not until its objectives have been reached should other subjects be attacked. In Wisconsin it has been the practice in most instances to set the dates for successive cooperative tests so that a period of two months or more may intervene between tests in different subjects. Arithmetic may be tested in November and reading in February. This allows the teachers and principals some time to work out remedial measures. If we may be permitted to offer a criticism upon our own work, it is that during the first year we failed to allow sufficient opportunity for this sort of concentration. In the second place, simple tests should be selected for use in all schools that have had no previous experience with tests, or experience of a very limited kind. In such schools, the many administrative details connected with the more elaborate tests only serve to obscure their real purpose.

AFTER THE WORK HAS BEEN WELL STARTED

The program outlined thus far has to do with the initial building of a foundation for educational research. After a group of superintendents has been trained to proceed with research and experimentation, the work of the director becomes somewhat different in character. If possible, an annual conference of school heads should be arranged. A number of states now hold such conferences, the best known being those held at Indiana University. At such gatherings the results of previous efforts can be reviewed. Matters for cooperative investigation proposed for the year can be discussed and voted upon. At such times it is the business of the director to bring to the attention of the assembled school heads the most recent developments in scientific procedure in education, and to call their attention to the best recent literature on the subject. If possible at this time, it will be well to bring before the conference some prominent workers in the field of educational research from the outside, who are capable of stimulating further educational investigation and professional growth. After a few days of wholesome discussion the director's suggestions and those made by superintendents and others can be finally considered and lines of investigation

voted upon. Superintendents should be made to feel that the choice is their own. Dates that are opportune should be chosen.

A bureau of research should at all times act as a source of information for teachers, supervisors, and superintendents on questions relating to the most effective methods in teaching, supervision, or administration. A part of its function is to suggest the most recent studies of practical value, to issue reviews of stimulating and suggestive articles and monographs in the field of education, and to send out bibliographies of the best publications pertaining to problems under investigation.

EDUCATIONAL MEASUREMENT IN WISCONSIN

Organized work in educational measurements in the schools of the state began in the fall of 1916 with the appointment of a director of educational measurements by State Superintendent C. P. Cary. Previous to this, the efforts to carry on educational measurements in the state were quite meager and more or less spasmodic. Since then a definite policy of cooperative educational investigation of school conditions has been pursued. Superintendent Cary's main purpose in establishing the bureau of educational measurements was to bring about a general improvement in the achievements of pupils throughout the state. It was his belief that instruction lacked the quality of thoroughness which should be demanded. The results of investigations to date have fully substantiated this position.

While attempting to carry out the purpose just mentioned, the bureau does not limit itself to efforts at improving the work of classroom teachers and supervisors directly. It seeks to improve administrative procedure and educational conditions. General school surveys and surveys of particular phases of local school problems are undertaken. The Janesville survey has been published as an example of a general survey. Among other things, efforts are made to secure increased financial support for schools, more economical distribution of school funds, better buildings and equipment, higher salaries for teachers, selection of better trained teachers, and improved gradation of pupils.

With the exception of the tests applied in cities during the first year, standardized educational tests have been given chiefly on a cooperative basis, made possible through the generous

response of two bodies—one the city superintendents and the other the county superintendents and supervising teachers. Each group is met annually in conference. At these conferences, plans are taken up for cooperative testing for the year. In addition to the groups mentioned above certain other groups, such as the teachers in state normal schools and those in the county training schools for teachers, are called upon at times.

The schedule carried out since the bureau has been in operation, the classes and numbers of schools reporting complete results, and the numbers of children tested may be seen in the accompanying table.

The foregoing tabulation by no means represents all of the testing done during this time. It usually happens that a number of schools find it inadvisable or inconvenient to give the tests at the particular time set, and that others make incomplete or late returns, or make no returns at all. Many schools give a number of tests in addition to the cooperative tests, depending upon the nature of the local problem. With the aid of the bureau a number have been making use of group intelligence tests. Superintendents are becoming more and more alert to the necessity of more accurate classification and gradation of pupils. During the year 1918-1919, the bureau administered its own group intelligence tests to students in seven of the state normal schools, as a means of aiding the normal schools in studying the quality of the prospective teaching material.

In addition to the testing, an increasingly large number of schools are engaging in definite experimental studies of special phases of instruction. During the current year intensive remedial work in reading is being emphasized more than any other topic.

The state supervisors, in their visits to schools, are using the facts discovered through the application of tests. They meet teachers, principals, and supervisors individually and in groups. At these meetings they discuss the test results and their applications to teaching method, course of study, equipment, and gradation of pupils. This personal contact of the state supervisors with the teachers—a contact in which the teachers are helped to make the applications—is an essential and vital part of the work as it is carried on from the Wisconsin State Department of Public Instruction.

SUMMARY OF COOPERATIVE EDUCATIONAL TESTING IN WISCONSIN
1916-1917¹

| Schools | Subjects | Test | Scope | No. of Children |
|---------------------|-------------|------------------|--------------|-----------------|
| Rural..... | Spelling | Ayres | 1173 schools | 15,825 |
| " | Writing | Thorndike | 173 " | 2,625 |
| "State graded" | | | | |
| (small village).... | Spelling | Ayres | 132 " | 7,465 |
| "State graded" | Writing | Thorndike | 9 " | 720 |
| City..... | Spelling | Ayres | 35 cities | 13,111 |
| " | Arithmetic | Woody-A | 21 " | 7,500 |
| " | Reading | Kansas | 18 " | 7,549 |
| " | Writing | Thorndike | 11 " | 3,866 |
| " | Composition | Hillegas, Trabue | 15 " | 5,848 |

1917-1918

| | | | | |
|------------|------------|-------------------|-------------|--------|
| Rural..... | Arithmetic | Stone-Original | 27 counties | 3,395 |
| " | Arithmetic | Woody-A | 31 " | 10,840 |
| City..... | Reading | Thorndike Alpha 2 | 34 cities | 9,821 |
| " | Arithmetic | Stone-Original | 59 " | 18,495 |
| " | " | " -Equivalent | 33 " | 11,010 |
| " | Algebra | Hotz | 22 " | 1,663 |

1918-1919

| | | | | | |
|--|------------|---------------|-------------|------------|-------|
| Rural..... | Spelling | Buckingham | 45 counties | 9,322 | |
| " | Reading | Monroe | 43 " | 9,090 | |
| City..... | " | " | 32 cities | 20,254 | |
| " | Arithmetic | Theisen-Woody | 29 " | 12,839 | |
| " | Latin | Brown | 15 " | | |
| Co. training and teacher training classes in H. S. . | { | Reading | Thorndike A | 50 schools | 1,334 |
| | | " | Rate | 41 " | |
| | | Arithmetic | Woody | 48 " | 1,304 |
| | | " | Stone | 44 " | 1,214 |

1919-1920

| | | | | |
|------------|------------|----------------|-------------------------|--|
| City..... | Spelling | Ashbaugh-Iowa | 114 ² cities | |
| " | Reading | Monroe Form II | ² | |
| " | History | Van Wagenen | ³ | |
| Rural..... | Arithmetic | Theisen-Woody | ² | |

¹ During the first year, 1916-1917, the tests in cities were given at the time of the director's visit. In succeeding years a definite date was fixed for the giving of the tests in all schools.

² Returns are just now coming in.

³ To be given during the present year.

Some attention is given to the training of prospective teachers in the use and application of standardized educational tests. Teachers should possess at least a preliminary understanding and appreciation of the purposes and values of such tests when they enter upon their duties. State normal schools are urged to offer courses giving specific training in the application of standardized tests. The state supervisors of educational measurement assist in the preparation of such courses. A number of Wisconsin state normal schools now offer training in educational measurements either as a separate course or as parts of other courses in education.

The personnel of the bureau includes, in addition to the director, an assistant director, Mrs. Cecile White Flemming, Supervisor of Educational Measurements, and a statistical secretary, Miss Elizabeth A. Ahern. Besides the immediate staff the bureau has, as has been said, the benefit of the cooperation and assistance of other members of the staff of the State Department of Public Instruction. Temporary clerical and statistical assistance is secured from time to time to aid in busy seasons and when material requested by superintendents is to be prepared. It is to be regretted that sufficient funds are not provided for clerical assistance of a permanent sort. This need is sorely felt by reason of the fact that many schools look to the bureau as a clearing house for information on such matters as comparative salaries, comparative school costs, school buildings and equipment, comparative figures on standard tests, age-grade statistics, etc.

Aside from its direct service to Wisconsin schools, the bureau of the state department is called upon from time to time for suggestions for the application of educational measurements and the development of new bureaus in other states. No doubt other bureaus have had many similar requests. Such requests are an indication of the growing demand for the application of scientific methods of procedure in educational administration and supervision.

PUTTING STANDARDIZED TESTS AND SCALES TO PRACTICAL USE IN THE RURAL SCHOOLS¹

SAMUEL S. BROOKS

District Superintendent, Silver Lake, New Hampshire

A supervisory district containing 26 rural, ungraded one-room schools with teachers for the most part inexperienced or untrained, a majority of whom had never worked under a superintendent before this, was the opportunity that fell to my lot when I became a New Hampshire superintendent under the new law effective September 1, 1919. It was my first experience as a superintendent. Truly the district was "virgin soil" as a member of the state department remarked to me when I took the position.

SIZING UP THE SITUATION

The first two or three weeks were spent in traveling about the district getting acquainted with the teachers and sizing up the situation. Most of the schools had already begun, the teachers having been hired by the local school boards as in the past. I found that three of the teachers were normal-school graduates, and that two had attended one summer school; but the others had no professional training whatever. Several had not attended school beyond the eighth grade, but had taught from ten to 28 years in the same school. Three were inexperienced high-school graduates in their 'teens. Fifty percent of all the teachers had never seen a professional book or magazine and did not know where to obtain one. All this of course had resulted in an ingrowing provincialism which could not but have a disastrously narrowing effect on their teaching. They were imitators of imitators. Their methods were in imitation of the teachers who had taught them and who, in their turn, had imitated their own teachers. The results were the use of methods and texts so archaic as to be amusing if they had not been at the same time such a sad commentary on our boasted educational system.

¹ This is the first of a series of articles describing some of the problems confronting an inexperienced superintendent of a rural district in New Hampshire and his attempts to solve them with the help of standardized educational tests and measurements.
—EDITOR.

Most of the textbooks were sadly out of date. Arithmetics, geographies, grammars, and even histories published in the eighties were in daily use. One of the local committees insisted that they must be good books or they would not have worn so long. Of course the books were all based on the defunct doctrine of formal discipline and showed little of the psychological methods of presenting material which are the basis of modern educational procedure. The idea seemed to be that an arithmetic is an arithmetic, and a geography is a geography and that one book is as good as another so long as nothing pertaining to the subject is omitted and the leaves are all present.

I have emphasized the teaching and textbook situations because I consider them the most important factors. A few other points I shall mention briefly. As for the schoolhouses, it is sufficient to say that most of them were typical New England rural school buildings of ancient vintage, modelled variously (according to the idea of their *instigators*), upon churches, town halls, barns, and in some cases it would almost seem, upon pigpens. They were small, dirty, poorly heated, lighted and ventilated, and in short generally unsanitary and ill-suited to their purpose. As to the organization of the schools, as I have already intimated, no attempt had ever been made to grade them. Another important factor in the situation was the attitude of the communities toward expert supervision. Three of the four towns in the district had always opposed such supervision as long as it was optional and were not inclined to submit gracefully when according to the provisions of the new law it became compulsory. Such were the general conditions as revealed by my preliminary survey.

However, there were three bright spots in the general darkness of the situation. First, the splendid new state school law which made the superintendent a state official with a pretty free hand and which provided liberal state aid for rural schools where it might be necessary to keep them up to the required minimum standards. Secondly, intelligent school boards willing to co-operate. And last, my own professional equipment which included a pretty thorough training in educational measurements and an earlier experience in the teaching of rural schools. Accordingly, I tackled the proposition determined to show what scientific method in education could do for rural schools if it had the chance.

ANALYZING THE SITUATION

The following week was spent in analyzing the conditions found to exist and in deciding upon the most efficient way to bring order out of chaos. As a result of this analysis a number of definite problems seemed to stand out clearly demanding early attention. From among these problems I shall select, for present discussion, only those, in the solution of which as it seemed to me, standardized tests and scales could be used to good advantage.

Problem 1.—To grade the schools fairly and accurately.—Getting the schools graded in order to start the pupils right for their year's work was of course the most urgent problem. It would have been a comparatively simple matter to grade them arbitrarily on the basis of the teachers' judgments. But there are serious objections to this method even as a beginning, especially when the teacher is new to the school. If, for example, children are placed in grades lower than their parents think they ought to be, there is bound to be some strenuous objections from some of these parents and the only evidence one has to offer is that the teacher thought the children ought to be placed in those grades. Did you never hear this line of talk from an irate parent? "My John is just as smart as Mrs. Smith's Henry and he ought to be in the same class or above. It's a plain case of showing partiality so there now! If just won't stand it!"

It is not by any means easy in such cases to prove to the satisfaction of the parent that the teacher has a sound basis for her judgment. And the fact is that she did not have a sound basis. How did she know just what knowledge or ability a child should exhibit in order to belong in a certain grade? What standards of achievement did she have for the different grades besides her own arbitrary judgment?

Giving ordinary tests for grading purposes is also inaccurate and unfair, because such tests are devised by the teacher or superintendent and hence represent merely their judgment as to what they think a child ought to know in order to be placed in a certain grade.

Worse than all else, the children themselves are apt to become discouraged, not understanding wherein they have failed, and being at a loss as to just what they are expected to achieve. Even the superintendent and teachers, if they are conscientious, are

not satisfied because they cannot feel sure that injustice has not been done. Retarding a child without good reason, thus robbing him of a year of his time, is a serious matter, at least for the child.

What method of grading then could be found which would be fair to all concerned? After casting about for some time for a solution to this problem, it suddenly occurred to me that standardized tests would help me out of my difficulty. Did they not offer definite standards of achievement for each grade? And those standards were not based on unattainable 100 percent perfection nor upon the opinions of dogmatic educators as to what a child ought to know at the end of a certain grade. They were experimentally derived and based upon the amount of work that normal children are actually doing in the different grades throughout the whole country and not upon the amount of work somebody thinks they ought to do.

By giving these tests in all my schools I could determine just where each child belonged on the educational ladder. The process would be roughly comparable to measuring a large number of sticks of various lengths by means of a tape measure sorting them into eight piles, each pile containing sticks of about the same length. Then when angry parents wanted to know why their children were graded so low, I could show them. The children themselves could be made to see wherein they had fallen below standard and what they would have to do to come up to the standard. Both superintendent and teachers would feel that they were on solid ground. Surely I needed standardized tests the first thing.

Having reached this conclusion, I immediately ordered all the tests and scales I knew of that were sufficiently well standardized for my purpose. They covered the following subjects: arithmetic, silent and oral reading, spelling, writing, geography, history and English language.²

Problem 2.—How to measure the progress of pupils.—Having settled upon how to do the preliminary grading, the next problem

² For those who do not know standardized tests and want to get acquainted I recommend for a beginning the purchase of *Measuring the Results of Teaching* by Monroe, and *Educational Tests and Measurements* by Monroe, DeVoss, and Kelly. Both books describe the best tests and their uses and the former tells where to obtain them. These books are published by Houghton Mifflin Company, Boston and elsewhere.

was to find a just method of measuring the progress of pupils for promotion purposes. The same objections to teachers' tests hold good here as were noted in discussing the previous problem.

This problem involved, obviously, the finding of a satisfactory system of marking. It has been proved beyond doubt that ordinary teachers' marks are unfair, inaccurate, and generally unsatisfactory as a means of measuring progress or as a basis for promotion.

In measuring the progress of pupils, standardized tests must surely find their widest field of usefulness. Here the process can be compared to standing a ten-foot measuring rod beside a young tree when it is a foot high and noting its growth in height. When it has added another twelve inches it belongs in the two-foot class. Twelve inches more and it is promoted to the three-foot class, and so on. It is even possible to devise a satisfactory system of marking, for report card purposes, with the aid of such tests, as will be explained later on.

Some of the teachers thought that the new state program of studies for elementary schools would be sufficient guide for promotion purposes. This program is thoroughly practical and up-to-date. But it offers no definite, objective standards of achievement from grade to grade. Therefore it is not a satisfactory tool with which to measure the progress of pupils and to determine when they are ready for promotion. It is true that a printed program of studies specifies the subject-matter to be covered by each grade but it is left entirely to the judgment of the teacher as to when that subject-matter is covered satisfactorily. The ground may be gone over more or less thoroughly and with very unequal results according to the methods and arbitrary standards of the individual teacher. Hence, it is obvious that programs of study cannot take the place of standardized tests that are even fairly objective.

Problem 3.—How to measure the ability of teachers.—One of the greatest needs of superintendents is a method of measuring the ability of teachers that will be accurate and fair to all concerned. A method that will not leave the superintendent open to the charge of favoritism or poor judgment. A method that will satisfy a teacher who has been rated something less than "excellent" that her shortcomings are real and not hallucinations

in the mind of the superintendent due to his personal dislikes and prejudices.

What is wanted is concrete evidence of a teacher's ability or lack of ability that will allay carping criticism from whatever source. It is decidedly neither accurate nor fair to estimate a teacher's ability solely by observation made by the superintendent during his visits to the classroom. It is not fair because (a) such observations do not furnish a sound basis for judgment, (b) the superintendent's opinions are quite apt to be colored by personal prejudices toward an individual teacher or her methods, (c) classes often show at their worst in the presence of visitors, (d) even the teacher may fail to do herself justice under the critical eye of the superintendent. It is inaccurate for all the reasons noted above and because (a) some teachers do excellent work when the superintendent is present and shirk all the rest of the time, and (b) if such teachers do their own testing, even the results may be made falsely to appear satisfactory.

Provided a teacher is of good moral character with high ideals and a fairly pleasant personality, her further desirability as a teacher is measured by the results she gets as determined by the progress of the pupils when such progress is objectively measured. After all, it is results we are after, primarily. Hence, the standardized tests measure objectively both the progress of the pupils and the ability of the teacher at one and the same time.

Problem 4.—To find a practical method of supervising study and of teaching how to study.—This may seem like a large contract for the one-teacher rural school with its crowded curriculum, but I shall try to show that it can be successfully worked out by using proper methods of study supplemented by judicious use of the standardized educational tests and measurements.

This paper merely presents some of the problems confronting me in my new work together with some plans for their solution and reasons therefor. Following articles will describe in detail the working out of these plans and the results achieved.

COEFFICIENTS OF REGRESSION

LEONARD P. AYRES
Russell Sage Foundation

In the March and April numbers of the JOURNAL OF EDUCATIONAL RESEARCH there were published articles by the writer on shorter methods for computing the coefficient of correlation. The object of the present article is to describe a short method for computing the coefficients of regression and to comment briefly on the practical utility of that measure of relationship.

The coefficient of regression is a measure which tells how much change takes place in one of two series of paired values for each change of one unit in the other series. Suppose we have a set of figures giving the ages and grades of the children in a school system. Since the plan of the system is that children shall be promoted one grade each year, it is sure that there will be a high correlation between the two series.

We know, however, that in nearly all school systems many children fail to go forward at the rate of a grade each year and that the average progress is in marked degree less than the ideal. The coefficient of regression is the measure which will tell us what the average rate of progress is. It will tell what the change is in grades for each change in years or what part of a grade the average child tends to advance for each additional year of age. This is what is meant in this particular case by the statement that the coefficient of regression shows the change in one series for a unit of change in the other.

The formula for the coefficient of regression, as it has ordinarily been stated, is not easy to use or understand, but its difficulties are mostly easy to avoid. Its use will be illustrated by finding it for two correlations, one for a short list of paired values and the other for a table of distribution. The formula for finding the coefficient of correlation is the one explained by the writer in the two previous articles published in this magazine. It is as follows:

$$r = \frac{\text{Sum of products of subject and relative items} - \text{av. of subject items} \times \text{tot. of relative items}}{\sqrt{\left(\frac{\text{Sum of squares of sub. items} - \text{av.} \times \text{total of these items}}{\text{av.} \times \text{total of these items}} \right) \times \left(\frac{\text{Sum of squares of rel. items} - \text{av.} \times \text{total of these items}}{\text{av.} \times \text{total of these items}} \right)}}$$

Let us apply this formula to a short list of paired values representing the ages and grades of ten school children. Two of them began school four years ago, two three years ago, two two years ago, two last year, and two this year. They all entered at the age of six. One of each pair went forward one grade each year while the other spent two years in each grade. Their ages and grades, together with the computation of the coefficient of correlation, are shown in the following tabulation:

| | Grades (Sub.) | Ages (Rel.) | Sub. ² | Rel. ² | S × R |
|---------|------------------|----------------|-------------------|-------------------|-------|
| | 1 | 6 | 1 | 36 | 6 |
| | 1 | 6 | 1 | 36 | 6 |
| | 1 | 7 | 1 | 49 | 7 |
| | 2 | 7 | 4 | 49 | 14 |
| | 2 | 8 | 4 | 64 | 16 |
| | 2 | 9 | 4 | 81 | 18 |
| | 3 | 8 | 9 | 64 | 24 |
| | 3 | 10 | 9 | 100 | 30 |
| | 4 | 9 | 16 | 81 | 36 |
| | 5 | 10 | 25 | 100 | 50 |
| Total | 24 | 80 | 74 | 660 | 207 |
| Average | 2.4 | 8 | | | |

$$80 \times 2.4 = 192$$

$$24 \times 2.4 = 57.6$$

$$80 \times 8.0 = 640$$

$$207 - 192 = 15$$

$$74 - 57.6 = 17.4$$

$$660 - 640 = 20$$

$$r = \frac{15}{\sqrt{17.4 \times 20}} = \frac{15}{\sqrt{348}} = \frac{15}{18.65} = 0.80$$

In this example the coefficient of correlation is equal to the fraction $\frac{15}{\sqrt{17.4 \times 20}}$. The numerator (15) is derived from the

products of the items of the two series. The denominator is equal to the square root of the product of two numbers, which are 17.4 and 20. The first of these two numbers is derived from the subject series, giving the data for grades, and the second from the relative series, giving the data for ages.

The coefficients of regression are derived from this fraction and are $\frac{15}{17.4}$ and $\frac{15}{20}$. The first has as its denominator the number

17.4, which was derived from the data as to grades. The coefficient of regression $\left[\frac{15}{17.4} \right]$ tells the amount of change in ages taking place on the average in this group for a change of one in the grades. This information does not in this particular case give us information of much value.

The other coefficient of regression is the fraction $\frac{15}{20}$ in which the denominator was derived from the data on ages. This coefficient tells us how much the children in this group go forward in the grades on the average for each advancing year of age. It tells us that they go forward $\frac{15}{20}$ or $\frac{3}{4}$ of a grade for each year of age. This is in accord with the original conditions of the problem, for it was stated that half the children made 100 percent of normal progress and the other half 50 percent, or an average of 75 percent of the full rate.

This is information of an important type that cannot be secured by the simpler processes of working directly with the averages of the series and trying to compute the relationship between the time the children have been in school and the progress they have made. In many sorts of results from tests of classroom work the coefficients of regression can be used to great advantage.

The method described for finding these coefficients is far simpler than the one usually used. The formula commonly given (rho standing for the coefficient of regression) is $\rho = r \frac{\sigma_1}{\sigma_2}$.

The steps by which the method here described was derived from the standard formula may be readily followed through by anyone familiar with the ordinary notations for the coefficient of correlation. They are as follows:

$$\rho = r \frac{\sigma_1}{\sigma_2}, \quad \text{or}$$

$$\rho = \frac{\sum xy}{n \sigma_1 \sigma_2} \times \frac{\sigma_1}{\sigma_2}, \quad \text{or}$$

$$\rho = \frac{\sum xy}{n \sqrt{\frac{\sum x^2}{n}} \sqrt{\frac{\sum y^2}{n}}} \times \frac{\sqrt{\frac{\sum x^2}{n}}}{\sqrt{\frac{\sum y^2}{n}}}, \quad \text{or}$$

$$\rho = \frac{\sum xy}{\sum y^2}$$

The application of the method described may be again illustrated by using as an example an age and grade table. The one chosen is a supposititious distribution for a system where 160 children enter each year at the age of six. They are promoted annually at a 75 percent rate and none drop out. The data for the table, together with the computations of the coefficients of correlation and regression, are shown in the accompanying tabulation.

| S | R—AGES | | | | | | | | T | ST | SST |
|--------|--------|-------|--------|--------|--------|--------|--------|--------|---------|------------------------|--------|
| Grades | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | | | |
| 1 | 160 | 40 | 10 | 2 | | | | | 212 | 212 | 212 |
| 2 | | 120 | 50 | 23 | 8 | 2 | | | 213 | 426 | 852 |
| 3 | | | 90 | 67 | 34 | 15 | 6 | 2 | 214 | 642 | 1,926 |
| 4 | | | | 68 | 67 | 42 | 22 | 9 | 208 | 832 | 3,328 |
| 5 | | | | | 51 | 63 | 47 | 29 | 190 | 950 | 4,750 |
| 6 | | | | | | 38 | 56 | 49 | 143 | 858 | 5,148 |
| 7 | | | | | | | 29 | 49 | 78 | 546 | 3,822 |
| 8 | | | | | | | | 22 | 22 | 176 | 1,408 |
| T | 160 | 160 | 160 | 160 | 160 | 160 | 160 | 160 | 1,280 | 4,642 | 21,446 |
| | | | | | | | | | | av. of sub. items | |
| RT | 960 | 1,120 | 1,280 | 1,440 | 1,600 | 1,760 | 1,920 | 2,080 | 12,160 | $\frac{4,642}{1,280}$ | =3.627 |
| RRT | 5,760 | 7,840 | 10,240 | 12,960 | 16,000 | 19,360 | 23,040 | 27,040 | 122,240 | | |
| ΣSf | 160 | 280 | 400 | 521 | 641 | 760 | 880 | 1,000 | | av. of rel. items | |
| R(ΣSf) | 960 | 1,960 | 3,200 | 4,689 | 6,410 | 8,360 | 10,560 | 13,000 | 49,139 | $\frac{12,160}{1,280}$ | =9.5 |

$$4,642 \times 9.5 = 44,099$$

$$4,642 \times 3.627 = 16,836$$

$$12,160 \times 9.5 = 115,520$$

$$49,139 - 44,099 = 5,040$$

$$21,446 - 16,836 = 4,610$$

$$122,240 - 115,520 = 6,720$$

$$r = \frac{5,040}{\sqrt{4,610 \times 6,720}} = \frac{5,040}{\sqrt{30,979,200}} = \frac{5,040}{5,566} = 0.91$$

$$\rho_1 = 1.09 \quad \rho_2 = 0.75$$

In this case the coefficient of correlation is 0.91. Such coefficients based on age and grade tables are always high and there is surprisingly small range between those for relatively efficient school systems, having little retardation and those for far less effective systems showing high percentages of retardation.

The coefficients of regression are of greater interest and more value. In this case the coefficient of regression based on the grades is 1.09. This indicates that in this school system the children tend to average 1.09 increase in years of age for each advance of one grade.

The other coefficient in this case gives us more valuable information. It is 0.75 and is based on the ages. It tells us that the children in this system tend to go forward on the average three-fourths of a grade for each advance of one year in age. This is in agreement with the conditions laid down at the outset which were that these children were regularly promoted at a 75 percent rate.

Coefficients of regression can be made to yield information of great practical value. They may well be used in studying such questions as: regularity of attendance and promotion rates, deficiency in weight and hemoglobin counts, progress through the grades and expenditures for supervision, pages of material read in the first three grades and ability in silent reading, time spent on spelling drills and scores in spelling ability, years of schooling and earnings up to the age of fifty, and many others of the most varied character.

Editorials

REVISING THE COURSE OF STUDY

In the voluminous discussion of the reconstruction of the curriculum during the past year nothing is said about the cost. It is assumed that if the superintendent and teachers decide upon a revision, the staff will perform the task in their leisure time. But such an assumption is entirely erroneous. A new course of study can be prepared by the staff but a thorough-going revision involves too much detail work of a clerical, time-consuming sort for any staff to be able to perform it in their free time. To discover the common errors in language in a school system as a basis for a language curriculum requires anywhere from two hundred to five hundred hours of clerical work, which is the full time of one person for at least a month of eight-hour days; and this does not include the time of the supervisor of the study who formulates the plans and writes the report. To discover what mathematics is needed as a preparation for the study of one textbook in chemistry involves fifty hours of clerical work in addition to forty hours devoted to the study by the supervisor. To collect practical problems and projects to supplement the teaching of physiology and hygiene, arrange the necessary material, and prepare it for classroom use involves the expenditure of fifty hours each for three persons. And all of these are fragmentary studies.

It is quite as reasonable to expect the manual arts department to build the school buildings without assistance as to expect the staff to revise the course of study in their free time. Bricks cannot be handled as rapidly as ideas but there are more ideas in the course of study than bricks in a building and it takes time to handle ideas. The task of the revision of what children are taught is immeasurably more important than the construction of buildings and the difficulties are much more complex and quite as time-consuming if the work is thoroughly done.

It is, therefore, necessary for boards of education who authorize a revision of the course of study to provide funds to make it

possible for certain teachers to be relieved of part of their duties, to engage clerical assistance, to employ experts on technical phases of the problem, and to make the necessary studies in the research department. It is becoming increasingly necessary for state departments, the federal bureau of education, and other large educational units to establish the curriculum as one of the major projects of investigation with a staff sufficiently large to make revision more than a makeshift. The deadweight of the load is much greater than administrators have generally realized. The work cannot be done by discussion and argument. Detailed studies are necessary and details cost money.

W. W. C.

SCIENTIFIC CURRICULUM BUILDING

Scientific curriculum building involves a danger of which the curriculum makers do not appear to be aware. The method as frequently practised consists in finding out what elements of school subjects are in frequent use and incorporating these elements into the subject matter to be taught. Analysis reveals, for example, that adults make almost no use of fractions or decimals; that periodicals and newspapers contain surprisingly few historical or geographic allusions; and that the fundamental vocabulary of business and personal correspondence is astonishingly small.

Sometimes the analyst reaches the conclusion, expressed or tacit, that the curriculum should contain *no more than* these meager elements which he has found to be in frequent use. More often, however, it is those who use his results who reach this conclusion. This, we maintain, is not a correct inference. The only correct inference—and it is the inference of the more careful workers—is that, no matter what else is included, the elements thus found to be socially useful shall certainly form part of the curriculum; in other words, that they shall constitute a part of the “minimum essentials” of the subject to which they belong.

The bearing of analyses of social needs upon the curriculum may be most completely illustrated with reference to spelling. This is true not only because the social value of spelling is obvious but also because more analyses have been made of spelling usage than of any other. The principles, however, which we shall suggest have a general application to all the subjects which have been similarly studied.

In the first place, we cannot by any means be sure that the present determination of frequency of use is the true one. For example, in the list which purports to consist of the thousand commonest words, we do not find the names of our daily meals—neither *breakfast* nor *lunch*, *dinner* nor *supper*. Nor do we find the commonest articles of tableware, such as *plate*, *cup*, *knife*, *fork*, and *spoon*. Many articles of food—*potatoes*, *beans*, *beets*, *peas*, *meat*, *fish*, *bread*, and *butter*—are not named; nor are some of the commonest articles of clothing nor some of the most familiar colors. Only the horse is mentioned among domestic animals. The dog and the cat, the cow and the sheep find no place in the list.

One may contend that these may indeed be common words without being frequently used in written discourse. The contention, however, has little plausibility. A child cannot write of his play without using such words as *hide*, *jump*, *swim*, *fish*, or *sled*, and these words are omitted from the thousand most frequently used. He can scarcely write of the barnyard without requiring such omitted words as *hen* and *egg*; he cannot write about his *garden* nor the *baby* at home; nor about the *moon*, the *sun*, the *sky* nor the *stars*. Yet these are the things within his experience. They are, therefore, the sort of topics which we should expect him to write about.

On the other hand, among the "thousand commonest words" are found these: *preliminary*, *recommend*, *impossible*, *assure*, *committee*, *investigate*, *national*; and a number of terms of apparently legal meaning such as *property*, *judgment*, *testimony*, *complaint*, *agreement*, *witness*, and *conviction*. This suggests that the source of our vocabulary studies must be very greatly widened and that an enormously greater amount of written material must be examined.

In the second place, even if true frequency of use has been determined, we assert that there are other tests of value than frequency of use. For example, although the word *bread* may be rightly excluded from the thousand commonest words; yet circumstances which, while they may be relatively infrequent are nevertheless inevitable, will certainly make it at times indispensable in one's written vocabulary. Future research may conceivably confirm the present determination that *apple* and *banana*, *bright* and *beautiful*, *meat* and *drink*, likewise do not belong to the

thousand commonest words. But this is not a sufficient reason why they should be banished from our spelling curriculum. These are simple words and they stand for ideas common to the race. The ideas may not require expression quite often enough to secure for the words a place among those used with an assigned frequency; but when these ideas do require expression—and no one can escape the necessity for such expression—they become as “socially useful” as do their more frequently used competitors.

In the third place, in selecting material for the school curriculum, difficulty as well as frequency of use is a criterion. Indeed to a certain extent there is an inverse relationship between difficulty and frequency of use. Some elements of school subjects, precisely because they are frequently used, and hence often repeated, do not need to be taught at all. Such is the case, for example, with most of the two and three letter words in our language. No one who is in the least degree literate misspells *and* except by inadvertency—and correcting inadvertency is not a spelling problem. Some of the recent spellers in their attempts to adhere to the vocabulary studies have included a considerable group of words which do not constitute a spelling problem.

In the fourth place, whether with reference to spelling or any other subject of the curriculum, whatever the much-discussed aim of education may be it contemplates something more than the round of common experiences, something richer than the elemental facts of life. “A man’s reach should exceed his grasp”; and it is only by providing for more than is now generally available that progress is possible. Here as elsewhere in the application of scientific procedure to education, a sound philosophy—as well as a sound common sense—must be invoked to save the scientific procedure from itself.

Four points may, therefore, be made with reference to the sufficiency of the idea of frequency of use. The first is that material far greater in amount and variety will have to be examined before we can be sure that frequency of use has been determined. The second is that certain material of relatively low frequency of occurrence has a social usefulness which renders it indispensable. The third is that difficulty, or need of being taught, is an important criterion—one which operates in particular to reject from the formal curriculum an appreciable amount of frequently used material. The fourth is that if we include in our curriculum only

the materials which are found to be frequently used by children and adults we leave altogether out of account the material which they ought to use and would use if their lives were fuller and richer. This fullness and richness, we feel quite sure, the school must envisage. Moreover, to adjust the curriculum merely to current usage is to doom our educational system to a static condition wherein progress becomes impossible.

B. R. B.

Reviews and Abstracts

VIRGINIA EDUCATION COMMISSION. *Virginia public schools*. (Including the report of the Survey Staff.) Richmond, Va.: Everett Waddey Co., 1919. 400 pp.

Very businesslike is this report. Its precision of aim gives it a certain directness and point which is often lacking in surveys whose object is merely to find out something. Moreover, the survey is peculiar in other respects. In the first place, it is a state survey and, while the system is not among the largest, it is by no means small. In the second place, considering the bigness of the undertaking and the wide range of topics discussed, the survey report is brief—brief in a day when brevity of reporting is by no means common. In the third place, being a state report, its findings work over inevitably into recommended legislation, i.e., into provisions for specific improvements. In the fourth place, there is an entire absence of either animus or the effort to “get” somebody.

Despite the brevity of the report, an abstract of it would be too extensive. Moreover, the book, we are told, will be published shortly by a commercial house and will, therefore, be easily available. We feel inclined strongly to recommend it to students, nor should we be greatly surprised if professors of education should find parts of it—especially the financial parts—useful as required reading.

As an example of a good method in reporting, it seems to us that the book is distinctive. The main body of tabular material is thrown into the back of the volume; this device adds to the readability of the report. We beg, however, to suggest an added improvement, namely, that the reading matter should be complete in itself; in other words, that the reader should be conscious of no loss of essential detail. Thus, the text should not directly require the reader to consult the tables. It should form a consecutive whole, and the references to tables, like bibliographical citations, should be carried in footnotes.

The legal status of schools in Virginia is peculiar in that the constitution, instead of confining itself to general provisions, embodies many details. Accordingly, many of the survey recommendations require constitutional amendments. Some of these amendments were proposed by the 1918 assembly; others will have to come up for the first time this year. They must then be passed again in 1922 and submitted to the people.

Some of the important recommendations of the commission to the General Assembly—recommendations in almost complete agreement with the corresponding recommendations of the survey staff to the commission—are as follows: (1) that the state board of education be appointed by the governor, and that it select the state superintendent at large; (2) that division superintendents be elected by local school boards from a list of eligibles prepared by the state board of education; (3) that the district unit be abolished in favor of the county (city or town) unit; (4) that the state school fund be distributed as follows: an amount not to exceed 5 percent to be used as a special relief fund; of the remainder, one-third to be distributed on average or aggregate attendance, one-third on the number of teachers employed, and one-third according to the adequacy of the local support; (5) that an effective compulsory education law and a school term of nine months be provided; (6) that assessments be increased and equalized and the constitutional tax limit removed, the limit to be determined by the legislature; (7) that free textbooks be provided at state expense; (8) that teachers' salaries be increased and based on minimum professional qualifications, the state tax being increased from fourteen to twenty cents to provide a part of the necessary funds; (9) that, as to teacher-training, normal schools abolish high-school courses, train only elementary school teachers, and extend the professional course from two to three years; that training departments in high schools be abandoned; that state institutions of higher learning be operated all the year round on the four-quarter basis, with special provision by means of two six-week terms in the summer for public school teachers; (10) that vocational education, physical training and school hygiene, education of negroes, and the consolidation and improvement of rural schools be definitely organized.

The survey has not failed to provide for adequate financial support for the public school system as it would be constituted if these recommendations were enacted into law. We should not know where to turn to find a clearer statement of how to determine and provide the amount of money needed to operate a good school system.

The General Assembly of 1918 had appropriated \$10,000 for the survey commission. Later it seemed advisable to create a division of tests and measurements. It was evident, however, that, in order to do anything at all adequate in applying educational and psychological tests throughout the state, a great deal more money would be needed than the survey had in hand. Fortunately, the General Education Board was induced to appropriate \$12,500 for this purpose.

Accordingly, for the first time in a state survey an attempt was made to ascertain the character of the work of the schools by means of actual standardized tests.

The chapter devoted to the work of this division is entitled: "The Results of Instruction Measured." It is wholly inadequate; and the propriety of including it in the survey volume is doubtful, especially since a special report on this phase of the survey work is promised for later publication. We observe the rather curious, not to say crotchety, selection of tests. This display of originality, however, led to at least one fortunate result, namely, the inclusion of two intelligence tests in the list. This is the first large use in a survey of such tests. The results of these tests, if, in the forthcoming publication, they are properly related to the educational tests, will constitute an important improvement upon conventional methods of reporting.

It seems to us that the work of the Division of Tests and Measurements forms no integral part of the survey which the Assembly had in mind when it created the commission. The problem was distinctly a legislative one. What Virginia needs more

than anything else is permission to go ahead and mandatory laws which will cut away old prejudices. The decision to test the schools may have been a wise one from another point of view, but it seems to us to constitute an almost entirely separate project.

As to the rest of the report, it is, as we have intimated, unusually effective. It "gets across." Unless we are much mistaken, a large proportion of the commission's program will be enacted into law—this, of course, unless politics play an unusually powerful rôle. The price of \$10,000 may well prove infinitesimal compared with the advantages.

B. R. B.

MANUEL, HERSCHEL T. *Talent in drawing*. (School and Home Education Monographs, No. 3.) Bloomington, Illinois: Public School Publishing Company, 1919. 152 pp.

This monograph reports a study of gifted children made in 1916-1917. It was subsidized by the General Education Board and was conducted under the direction of Professor G. M. Whipple. In this study two groups of important related problems were recognized, namely (1) those pertaining to pupils of high general ability, and (2) those pertaining to pupils of special ability. Dr. Manuel's book has to do with the second group of problems. Its scope is indicated by its subtitle: "An experimental study of the use of tests to discover special ability." Pupils of special ability are defined as those who "by endowment or by training are so equipped as to take special interest in, and to succeed extraordinarily well with, some one ability."

The present study proposes to consider a special ability (drawing) with reference to two problems:

1. What are the essential psychophysical characteristics of persons possessing talent in drawing?

2. How may the test method be used in the diagnosis of such talent?

Forty-five individual tests, mental and physical, which were more or less standardized, were used with nineteen subjects chosen for recognized talent in drawing. The subjects were distributed as follows: college or university, 5; secondary school, 8; elementary school, 6. Biographical and personal data were also collected. The test results are presented by means of the "psychograph" or "profile" showing graphically the performances of each subject in several tests as compared with standard performances. This method of representing the statistical data derived from psychological tests is especially useful for diagnostic purposes.

As to the first problem—the characteristics of those talented in drawing—the following abilities seem closely related to ability in drawing: Mentally to note a visual form and to reproduce it; to observe; to select from a complex visual situation the most representative and the most beautiful aspects; to remember visual forms; to manipulate them mentally; to control hand movements in accordance with visual percept or image; to invent—i.e., to bring together into new artistic combinations the elements of different visual experiences; to judge the beautiful in line, form, color, and composition; and to discriminate differences in color and in visual magnitude. Acuity of vision, interest in the act and products of drawing, and general intelligence are likewise characteristics of those talented in drawing.

As to the second problem—the diagnosis of talent in drawing—naturally tests of the foregoing abilities are the ones selected from the entire forty-five as having highest diagnostic value. These are described and references are given indicating

where they may be found. It is recommended that tests be arranged to ascertain not only the strength of an ability but also its improvability. The so-called "instructed-learning" method is therefore advised. Finally, the author notes that "the measurement of many different abilities not primarily related to drawing is necessary to the most effective educational and vocational use of measurement of ability in drawing."

While it has not been the purpose of the author in this study to prepare material and standards for the recommended tests, he promises a further contribution of this nature. His limitations due to lack of standards is evident throughout, but we trust the one who has made so thorough an investigation of this particular problem will shortly supply the supervisor and teacher of drawing with usable tests valuable in determining the particular talent.

J. ORIN POWERS

MURDOCH, KATHARINE. *The measurement of certain elements of hand sewing.* (Teachers College Contributions to Education, No. 103.) New York: Teachers College, Columbia University, 1919. 120 pp.

The scope of this monograph is well described by the following statement of the author:

"The specific tasks of this research are as follows:

1. To make a scale by means of which merit in certain forms of hand sewing may be measured.
2. To make an inventory and analysis of the faults found in children's sewing.
3. To determine the relative importance of various faults and of various stitches as indicators of general merit in sewing.
4. To determine the relative reliability of judgments concerning various faults and concerning various stitches.
5. To determine (a) the reliability of one or more samples of a child's sewing in evaluating her real ability in that kind of sewing; to determine (b) the number of samples of various stitches which would be required in order to have various degrees of probability that a child's real ability to execute that particular stitch was measured.
6. To determine the relative accuracy of competent persons, with varying degrees of knowledge about sewing and the teaching of sewing, as judges of sewing products.
7. To test the value of our sewing scale as a means of attaining greater exactitude in the grading of sewing products."

As a preliminary step in constructing the scale for measuring certain elements of hand sewing, samples of the hand sewing of 1,212 individuals were secured. These samples are called samplers. They were uniform as to material and kinds of stitches. In quality they ranged from the samplers contributed by 37 children who were in special classes for defectives or in institutions for the feeble-minded to samplers contributed by adults who were judged to be expert hand sewers. From these samples of hand sewing 15 samplers were finally selected, which, on the basis of the judgments of competent judges, range in quality from 0 to 16.4. The method of selection and of determining the scale values follows closely that originally used by Thorndike in the field of handwriting and drawing. Each unit "represents that amount of merit which

exists between the samplers when one is recognized as being of more value than the other by 75 per cent of competent judges and of less value than by the other 25 per cent of the judges." "Zero was arbitrarily defined as meaning the score to be given a sampler which could be recognized as an attempt to conform to the directions for the making of the sampler, but which had absolutely no merit as an example of the stitches indicated." On the basis of this definition of zero, 38 of the 80 judges said that the sampler finally chosen to represent zero, represented zero, 21 judges considered it better than zero, and 21 judges considered it worse than zero. This sampler was contributed by an imbecile boy "untrained in sewing who is particularly deficient in motor control and who suffers from serious eye trouble which often causes dizziness and nausea."

Each of the samplers contains five kinds of stitches: combination, running, over-casting, hemming, and backstitch. Half-tone reproductions have been made of the fifteen samplers chosen for the final scale. Three views are given of each sampler in such a way that all of the sewing is represented. Thus the final scale as made available for general use consists of half-tone reproductions of the fifteen samplers made in such a way that all of the sewing is clearly shown.

The construction of the scale was supplemented by the analysis of the faults in the stitches involved in these samples of hand sewing. The list of faults was finally reduced to 23 in number and the relation of these to general merit was studied. This work supplements the scale in a very significant way. It is particularly valuable in assisting in the interpretation of measures in terms of the pupil needs.

In the study of the reliability of the measures yielded by the scale there are two significant findings. When it was used to measure the quality of the hand sewing in completed articles such as aprons and bags, it was found that the measures assigned by different judges were less variable than when isolated hand sewing on samplers similar to those of which the scale was made was measured. This finding shows that it is feasible to use the scale in measuring the hand sewing done in making real things. The second finding was that judgments of different teachers, based upon the use of the scale, were found to be less variable and hence more accurate than judgments made without the scale, although only teachers from a single large high school were used in this experiment and were untrained in the use of the scale.

The most significant features of Miss Murdoch's work are the ways in which she has supplemented the usual scale derivation. The time is rapidly approaching, if it has not already arrived, when new measuring instruments will receive little consideration until scientific evidence concerning their usefulness is made available. Already it is being urged by some that each measuring instrument should be accompanied by a definite statement concerning its reliability and validity. Miss Murdoch is one of the first, if not the first, to accompany the account of the derivation of a scale with a statement of the reliability of the scores that are obtained by using it. Furthermore, as stated above she demonstrates that the scale can be successfully used in judging the hand sewing on completed articles of a rather complex nature.

The scale has yet to be standardized and there is not contained in the monograph sufficient directions to guide a teacher in using it. It is to be hoped that in making the scale itself available to teachers, detailed directions for using it as well as record sheets for the tabulating of the results will be provided.

W. S. M.

HENRY, THEODORE S. *Classroom problems in the education of gifted children.* (Nineteenth Yearbook of the National Society for the Study of Education.) Bloomington, Illinois: Public School Publishing Co., 1920. 125 pp.

For a long time there has been in this country some recognition of the necessity of meeting the needs of pupils of marked ability in such a way as to give them opportunity for more rapid advancement than that of the average pupil. The great and hitherto unsuspected individual differences revealed by the recent widespread application of mental and educational tests are fast pushing this subject to the forefront of attention.

In this monograph the National Society for the Study of Education has once more made an important and timely contribution to the science of education. The study is the report of the results of an experiment in which pupils selected on the basis of their excellence in school work were placed under special instruction. It includes a summary of the various schemes for flexible promotion which have been tried out in this country, and a description of the types of special classes that have been established for gifted pupils. In addition, there is a chapter on the methods of teaching which are especially adapted to the instruction of this class of children, and a bibliography covering the general field.

The number of pupils selected for the experiment which was carried out in one of the public schools of Urbana, Illinois, was thirty—fifteen from the fifth and fifteen from the sixth grade. The attempt was made to carry out the experiment under average conditions of schoolroom instruction. "Supervision, course of study, physical equipment of room and of building, instruction—none of these could be said to be above average." The remaining fifth- and sixth-grade pupils of the school were used throughout the experiment as a control group. The intelligence quotients of the members of the special school as measured by the Terman revision of the Binet-Simon Scale ranged from 99.3 to 146.6. It appears, therefore, that on the basis of these tests there were pupils in the special class who did not belong there. Indeed, if the author's recommendation were adopted that a minimal I. Q. of 115 be the standard for admission to such special classes, it would rule out fourteen of the thirty pupils of this special class. Notwithstanding this handicap the special class was able to do the work of two grades in one year, and seventeen pupils secured corresponding promotion at the end of the school year on the basis of the regular examinations. Almost all of the remaining pupils passed special examinations after review work during the summer and thus secured promotion. Educational tests given throughout the year showed that by May first the pupils had actually gained a year as measured by these scales.

The interpretation of the results of such an experiment is necessarily somewhat difficult. There is no way of determining to what extent the more rapid progress of these pupils is due to the special conditions under which they were placed rather than to their superior abilities. As the author points out, practically all the methods adapted to the instruction of the specially gifted are equally applicable in the teaching of the average pupil. It is difficult to say how much more rapid the advancement of the control group might have been if it had had the advantage of the same special attention, and especially if the same spur to ambition had been provided. The success of a large number of pupils of only average intelligence in gaining promotion seems to show that these factors were more important than the author believes.

EDWARD H. CAMERON

PUBLICATIONS RECEIVED

- Alexander, Carter. *School statistics and publicity*. Boston: Silver, Burdett and Co., 1919. 332 pp.
- Ashley, Roscoe L. *Early European civilization*. New York: Macmillan Co., 1919. 714 pp.
- Ashley, Roscoe L. *The new civics. A textbook for secondary schools*. New York: Macmillan Co., 1918. 420 pp.
- Buckner, Chester A. *Educational diagnosis of individual pupils*. (Teachers College Contributions to Education, No. 98.) New York: Teachers College, Columbia University, 1919. 93 pp.
- Canby, Henry S. and Opdycke, John B. *Elements of composition for secondary schools*. New York: Macmillan Co., 1919. 593 pp.
- Crawford, Douglas G. *The study of English*. New York: Macmillan Co., 1919. 338 pp.
- Curtis, Francis D. *A manual of experiments in elementary science*. (Student's edition) New York: Chas. E. Merrill Co., 1918. 101 pp.
- Curtis, Francis D. *A manual of experiments in elementary science*. (Teacher's edition) New York: Chas. E. Merrill Co., 1918. 231 pp.
- Durell, Fletcher and Arnold, E. E. *A first book in algebra*. New York: Chas. E. Merrill Co., 1919. 364 pp.
- Dyer, Franklin B. and Brady, Mary J. *Readings in literature*. (Merrill readers) New York: Chas. E. Merrill Co., 1918. Books 1-2 for seventh and eighth grades.
- Dyer, Franklin B. and Brady, Mary J. *Suggestions to teachers. A manual to accompany the Merrill readers*. New York: Chas. E. Merrill Co., 1920. 209 pp.
- Lewis, Ida B. *The education of girls in China*. (Teachers College Contributions to Education, No. 104.) New York: Teachers College, Columbia University, 1919. 92 pp.
- Mackie, Ranson A. *Education during adolescence*. New York: E. P. Dutton and Co., 1920. 222 pp.
- Pargment, M. S. *Exercices Français*. New York: Macmillan Co., 1920. 2 parts.
- Russell, Frances T. *Satire in the Victorian novel*. New York: Macmillan Co., 1920. 335 pp.
- Scrimshaw, Stewart. *Bricklaying in modern practice*. New York: Macmillan Co., 1920. 182 pp.
- Searson, J. W., Martin, G. E., and Tinley, L. W. *Studies in reading*. (Teacher's manual) Lincoln, Nebraska: University Publishing Co., 1919. 319 pp.
- Seashore, Carl E. *The psychology of musical talent*. Boston: Silver, Burdett and Co., 1919. 288 pp.
- Smith, Wayne P. and Jewett, Edmund G. *An introduction to the study of science*. New York: Macmillan Co., 1919. 620 pp.
- Trafton, Gilbert H. *Science of home and community*. New York: Macmillan Co., 1920. 561 pp.
- Wallin, J. E. Wallace. *Problems of subnormality*. Yonkers, N. Y.: World Book Co., 1917. 485 pp.
- Winslow, Charles E. A. *Healthy living*. New York: Chas. E. Merrill Co., 1920. 2 v.

News Items and Communications

This department will contain news items regarding research workers and their activities. It will also serve as a clearing house for more formal communications on similar topics, preferably of not more than five hundred words. These communications will be printed over the signatures of the authors. Address all correspondence concerning this department to Walter S. Monroe, University of Illinois, Urbana, Illinois.

West Virginia University Bureau of Educational Research

One of the latest additions to the rapidly increasing group of state educational institutions which have established a bureau of educational research is West Virginia University. In a recent circular they announce the establishment of the bureau and give the prices of a number of educational tests and the addresses from which they can be obtained.

In 1917 Dr. Albert Shiels, then superintendent of schools, recommended the establishment of two new departments. Accordingly a Division of Educational Research was established in charge of Mr. Robert H. Lane, who as a principal of one of the elementary schools had attracted attention. Also, a Division of Psychology was created and Dr. A. H. Sutherland was invited to begin some studies on types of children. In 1920 Mr. Robert H. Lane was elevated to an assistant superintendency and thereby rewarded for meritorious service.

During the past two years comments have not been lacking that the work of the divisions overlaps. It was therefore decided to bring the two divisions together under the temporary name of the Department of Psychology and Educational Research with Dr. A. H. Sutherland as director. The new department is to continue the functions previously served by both divisions. It will have charge of the giving and training teachers to give standardized educational tests and the collection and tabulation of statistical information. It will also have charge of the segregation of types of children and the direction of the special rooms for backward children.

Mr. H. C. Daley, Survey Department, Highland Park, Michigan, sends the following tabulation of the results of giving Monroe's Standardized Silent Reading Test in the seventh and eighth grades. Form 3 was given on September 19, 1919, and Form 2 on January 28, 1920. "After the September testing pupils were reclassified into strong and weak sections in accordance with the results of the tests. In several cases capable pupils were transferred to a higher grade." This shifting would tend to reduce the standing of the groups. In spite of this, all groups show remarkable gains.

This use of two different forms of this test raises the question of the equivalence of the forms. Although the different forms were constructed so that they would yield

equivalent comprehension scores, a recent study of the different forms shows that this was not accomplished in all cases. However, the best data available at this time indicate that Forms 2 and 3 are equivalent for comprehension in the seventh and eighth grades. They are not equivalent with respect to rate. The rate scores on Form 2 may be expected to be fifteen to twenty words higher than on Form 3. Thus the non-equivalence of the two forms used probably accounts for a portion of the gain in rate but it does not account for any of the gain in comprehension.

| GRADE | No. OF CHILDREN | | RATE | | COMPREHENSION | |
|--------------------|-----------------|------|-------|------|---------------|------|
| | Sept. | Jan. | Sept. | Jan. | Sept. | Jan. |
| Low seventh | 252 | 205 | 102 | 118 | 21 | 27 |
| High seventh | 170 | 136 | 103 | 134 | 22 | 28 |
| Low eighth | 199 | 161 | 101 | 135 | 21 | 28 |
| High eighth | 156 | 285 | 104 | 124 | 24 | 28 |

| NUMBER OF PUPILS: | LOW SEVENTH | | HIGHSEVENTH | | LOW EIGHTH | | HIGH EIGHTH | |
|--------------------------------------|-------------|------|-------------|------|------------|------|-------------|------|
| | Sept. | Jan. | Sept. | Jan. | Sept. | Jan. | Sept. | Jan. |
| Above rate 130 words per minute..... | 39 | 96 | 22 | 72 | 23 | 93 | 29 | 95 |
| Below rate 75 words per minute..... | 29 | 14 | 29 | 2 | 36 | 4 | 16 | 16 |
| Above comprehension of 30..... | 35 | 76 | 27 | 53 | 42 | 68 | 48 | 115 |
| Below comprehension of 11..... | 44 | 2 | 14 | 1 | 29 | 3 | 9 | 9 |

The twenty-eighth annual meeting of the American Psychological Association, held at Harvard University, December 29, 30, and 31, 1919 was attended by more than 150 workers in psychology and allied fields. Joint sessions were held with the American Anthropological Association and the American Psychological Association of Clinical Psychologists and provisions were made for inspection of the Massachusetts State School for Feeble-Minded, the Judge Baker Foundation, the Carnegie Nutrition Laboratory, McLean Hospital, the Psychopathic Hospital and other local institutions.

Shepherd I. Franz, Government Hospital for the Insane, Washington, D. C., was elected president of the Association for 1920; H. S. Langfeld, Harvard University, and W. V. Bingham, Carnegie Institute of Technology, were elected to the council. Chicago was elected for the convention in December, 1920. Twenty-seven new members were voted into the association and the deaths, during the year, of John Wallace Baird, Edward Cowles, and August Hoch were reported.

A committee consisting of Bird T. Baldwin, University of Iowa (Chairman); W. F. Dearborn, Harvard University; L. S. Hollingworth, Teachers College; B. Ruml,

The Scott Co., Philadelphia; and Helen T. Wooley, Vocational Bureau, Public Schools, Cincinnati, Ohio, were appointed to consider the legal and professional status of "clinical psychologists." The N.E.A. at its last meeting passed a resolution to the effect that the diagnosis and segregation of defective children should be in the hands of highly trained and certified psychologists only; a similar resolution being adopted by the Educational Conference held at Albany, N. Y., in May 1919. The committee listed above is now engaged in (1) determining the qualifications of a clinical psychologist and in suggesting means of establishing and promoting an adequate professional curriculum; (2) investigating the methods of legal certification now operating in the states; (3) compiling data on the legal aspects of clinical work—the acceptance of testimony of a psychologist in court, commitments, etc.; (4) drawing up plans whereby the American Psychological Association may issue certificates to "clinical psychologists" and (5) determining the qualifications for psychological workers in industry.

The program of the session gave 6 titles under general psychology, experimental psychology 16, intelligence tests, educational and clinical psychology 21, comparative psychology 2, social and religious psychology 3, and applied psychology 7. The pronounced development of the technic of trade testing, the thorough tryout of intelligence tests by their use on more than a million and a half recruits, bringing up the question of revision of older notions of median mental levels of unselected as well as psychoneurotic, foreign-born, colored and other individuals, the need of tests other than those of general intelligence for diagnosis of psychoneurotics, vocational aptitudes and retardation in schools, the need of systematic training and classification of workers in the rapidly growing field of clinical psychology and the use of individual and group tests in diagnosis and prognosis of ability in school functions, may be considered the main features of the meetings.

The papers were so numerous and so varied in content, that no attempt will be made here to allude more than briefly to those which bear closely upon educational topics.

The address of the retiring president, Walter Dill Scott, Northwestern University, on the topic "Changes in Some of Our Conceptions and Practises in Personnel" gave a survey of progress in educational theory and practice. Evolution along five lines was emphasized: (1) A history was given of changes in the concept of the amount and nature of individual differences from the older belief that peoples were originally much alike and that such differences as appeared were due to environmental factors to the present belief that individuals differ enormously in all traits and that such differences are largely innate. "The greatest achievement of the association has been the establishment of the fact of individual differences and its applications to education, politics, industry, and economics."

(2) The importance of reasoning as a factor in adjustments has been superseded by the consideration of instincts, emotions, and habits. The operation of instinctive trends as the spring of action and the sources of discontent in industrial life was illustrated by concrete cases. It is not the violation of the worker's logical processes that brings unrest and discontent so much as thwartings of his pride, his desire of social approval, of mastery, and his instinctive and emotional trends.

(3) Education in the schools or in the factory must take into account the unlearned forms of reaction to features of the environment and seek to so order the individual

and his environment that the desirable adjustments may be made. The work of the personal director in industry must be enlarged and his function must be that of providing the approved experiences so that desirable habits of response will be built up. Mere richness of content or training of general faculties will not suffice.

(4) Changes have taken place in the concept of man and his environment. A man is not the victim of his environment, he is not the master of it nor is his function that of subduing or opposing it. The two should evolve together. In industry, the man and his job are not opposed, nor does the concept of selecting the man to fit the job—finding a square peg for a square hole—adequately express the present point of view. The function of the personnel manager is the “creation of the worker in a working situation.” The work must be arranged to satisfy the human wants of the worker and he should be set to the task most in harmony with his particular talents and interests. There must be a correlated improvement of the work and the workers.

(5) The last point led to a survey of the evolution of vocational guidance from the chance methods of superstition, guess, and incidental inclinations to the modern technic to mental, physical, and vocational tests, statistical methods and the like. “In the last century the productivity of the worker has been doubled by the activities of some 2,000 minds through changes in the material world. In the present century equally great increases in productivity may be expected through the adjustment of worker to the work.”

Following are brief summaries of papers, presented in the various sessions, which may be of interest to workers in education.

A most valuable and practical method of evaluating school attainments in terms of mental ability is the result of studies by Rudolf Pintner of Ohio State University. A mental test alone is not sufficient to diagnose a child accurately; measures of school attainment based on both age and grade norms are needed. Two survey tests, one mental and one educational, were devised and applied to 1,500 cases. Percentile values are converted into a so-called “index” ranging from zero to one hundred. The median indices serve as a measure for the school. A high mental age with low educational index indicates deficiencies in instruction, whereas a close agreement between the two indices indicates efficiency in the school. In the same way different classes in the same school or different individuals in the same class may be compared, enabling one to correlate innate ability with possible attainment and to apply corrective measures when the need appears.

E. A. Doll, Princeton University, presented data obtained from intelligence tests of a million and a half army recruits and other subjects bearing upon the median mental age of adults, commonly assumed to be about 16 years for unselected groups. The Stanford-Binet median mental age equivalent of army recruits is 13 years; of negro and foreign-born recruits 10 years; of adult male state prisoners 13 years, the actual tests used in each case being the Army Alpha. Five hundred typical public school children examined by Alpha show no increase in median scores by age after 13, the same being true of juvenile delinquents, ages 16 to 30 years. From these data the suggestion was made that the median mental age level of native white adults is approximately 13 years. No attempt was made to determine the life age limit of mental growth, since, it was suggested, “emotional development, skill, acquisition, aptitudes and the like probably continue to develop indefinitely.” The upper age limit of feeble-mindedness is not coincident with the lower mental age limit of nor-

mality, since it appeared that the borderline zone for feeble-mindedness may cover a range from mental age 7 to 13 years. The application of mental age or I. Q. as criterion of mental defect is specially limited in the case of individuals of 10 or more years chronological age. Data harmonizing with these results were presented by F. L. Wells of McLean Hospital. The median I. Q. (Stanford scale) of 102 cases of mental breakdown at McLean is 88 but I. Q.'s of 100 are frequent and I. Q.'s of 119 have been found in patients conspicuously incapable of self-maintenance. Only in organic psychoses does the breakdown regularly involve the ideational capacities with which the intelligence scales are concerned. Normal and even superior "intelligence" is very generally associated with grave judgment and conduct disorder. Intelligence scales measure essentially ability to deal with ideas, as distinct from ability to deal with things or with people. Psychotic breakdowns are essentially failures of adjustment to the social environment. The speaker emphasized the minor rôle of "intelligence" (ideational capacity) in mental balance, urging a conception of education as a discipline of character rather than knowledge.

With a few changes, the Binet tests were successfully employed by S. P. Hayes, Mt. Holyoke College, in measuring the intelligence of the blind. A modification of the Pressey group tests also gave high correlations with the Binet scores and abilities estimated by teachers. E. L. Woods, Wisconsin State Department of Public Instruction, reported on the use of the recently devised Virginia tests Alpha in group testing of intelligence of delinquent girls. J. B. Miner, Carnegie Institute of Technology, displayed a method of utilizing three-dimensional models for representing individual differences in complex abilities such as salesmanship. Certain relations not now taken care of in frequency tables, correlations, multiple regression, etc., can be displayed by such models.

L. W. Webb, Northwestern University, found the Pearson coefficient of correlation between rate of reading (Monroe Standardized Silent Reading Test) and the Army Alpha and Thurstone Tests A and B ranging from $+0.47$ to $+0.59$; comprehension of reading gave coefficients ranging from $+0.48$ to $+0.69$. Speed and comprehension of reading correlate $+0.85$. The speaker contended that these correlations show too large a dependence of the pencil and paper tests upon rapidity of comprehension in reading.

L. L. Thurstone, Carnegie Institute of Technology, found the cycle-omnibus form of intelligence test to be of great service in furnishing data for recommending candidates for admission to engineering colleges, in advising committees on scholarship in cases of delinquent students, and in giving vocational counsel. S. S. Colvin, of Brown University, reported a correlation of $+0.60$ between success in the Brown Intelligence Tests and the standing in college for the first two years. The Thorndike tests, first used in 1919, proved to be of great prognostic value and showed a high correlation with itself in repeated tests on the same subjects.

D. Mitchell, Pelman Institute, defined clinical psychology as a "professional practice" as contrasted with the "science" of laboratory psychology. Clinical psychology was again contrasted with "applied psychology" largely on grounds that the latter field may have no relation to the individual. Clinical psychology is the practise of determining mental states in order to prescribe kinds and methods of education, of detecting specific abilities and disabilities of vocational significance and the manipulation of emotional development for desirable social reactions. Dr. Augusta F.

Bronner, Judge Baker Foundation, emphasizing the function of clinical psychology as that of individual diagnosis, gave illustrations of the failure of general intelligence tests to betray significant variations in mental equipment. The necessity of devising tests for many specific abilities was urged. Florence Mateer, Bureau of Juvenile Research, has found that whereas the mental age or I. Q. are of little value in the detection of psychopathic conditions due to syphilis, a detailed account of the range of plus and minus scores, and more precise measures of the individual responses often give a reliable clue to either congenital or acquired syphilis. The same type of analysis generally differentiates the psychopathic from the non-psychopathic no matter what the mental age may be.

Tests devised to measure such functions as persistence, conscientiousness, application and emotional control are being applied to junior high-school and other students by S. L. Pressey of Indiana University. Correlations of results from these tests with general intelligence, health, and estimates of intelligence were presented. Mrs. L. W. Pressey, Indiana University, urged group testing of general intelligence in the primary school as a means of studying the effect of school training upon tests subsequently given. Data permitting comparisons of test findings on children of the laboring and professional classes, whites and negroes, etc., were presented. E. A. Kirkpatrick, Fitchburg State Normal School, believes that normal school pupils may be trained in fifteen hours to use the Binet tests in classifying pupils. Less time is required to handle group tests, but much more is needed to enable the student to score, tabulate and interpret the results. Daniel Starch, University of Wisconsin, found in data obtained from 1,000 first-year pupils, that approximately equal mental ability may exist with great differences in language ability. Correlations ranging from $+0.25$ to $+0.28$ between interest in college subjects and grades received in them were found by J. W. Bridges, Ohio State University. The correlation between interest and the subject's estimate of his ability in the courses was higher $+0.50$ to $+0.59$.

From data obtained in several extensive investigations cited by E. L. Thorndike, Teachers College, Columbia University, it appears that even experienced judges can not treat an individual as a compound of separate qualities, such as intelligence, industry, technical skill, reliability, etc., and to assign a magnitude to each of these in independence of the others. The correlations between such traits are too high and too even, the ratings being apparently affected by a marked tendency to think of the person in general as rather good or rather inferior and to be influenced in all judgments by this general attitude. It was suggested that observers should report *the evidence*, not a rating, and that the ratings should be given on the evidence for each quality separately.

E. G. Boring, of Clark University, offered several cautions concerning the uncritical acceptance of the distribution of mental traits according to the normal probability curve. The form of the distribution depends, for one thing, upon the unit of measurement adopted. If the normal law, moreover, is the law of chance, then two mutually dependent variables for which the relationship is not simply linear, such as height and volume which varies roughly as its cube, cannot have the same form of chance distribution. If the distribution for height is normal, the distribution for volume cannot possibly be.

W. T. Shepherd, Washington, D. C., in a study of 148 children, ages eight to thirteen years, emphasized the importance of the educational and social experiences, the influence of respected teachers or friends upon the development of religious ideals.

The prevalent belief that our judgments in the fields of appreciation of music or literature as well as in morals are considerably influenced by the contrary opinions of the majority of experts in these fields is supported by experimental data obtained from college students by H. T. Moore of Dartmouth College. Comparisons with control groups yield results showing that the knowledge of contrary opinions of the expert or the majority produces reversals of individual opinions in matter of speech amounting to four or five times mere chance. Moral reversals under majority influence are 4.7 times chance whereas reversals of musical preferences as a result of expert or majority opinion is about twice chance.

ARTHUR I. GATES

Teachers College, Columbia University

The State Department of Education of Colorado recently decided to have given in all schools of the state the following standardized tests:

A State-Wide Use of Educational Tests Courtis Standard Research Tests in Arithmetic, Series B, Gray's Oral Reading Test, Gray's Silent Reading Test, Ayres Spelling Scale, Ayres Handwriting Scale, "Gettysburg Edition."

At the present time the University of Colorado at Boulder, the State Teachers College at Greeley, and the State Normal School at Gunnison, are co-operating in conducting extension classes for Colorado teachers. For the western district of the state this work is in charge of Mr. O. B. Staples assisted by Dr. H. T. Manuel and Miss Stella Yowell. Mr. Staples writes as follows:

"The County Superintendents have been asked to direct this work (the testing program), and while they are struggling with it, those in my territory are quite unfamiliar with the technique of the work, and many of the teachers scarcely comprehend the significance of it, some of them not so much as ever having heard before of a "Standard Test." But this action of the State Department has given the work large publicity, and Dr. Manuel and I are cooperating with the people in our territory, meeting them in groups, singly, etc., giving them such instruction, encouragement, etc., as we can to help them "put over" the program of the State Department, although we think it is being attempted in a very crude manner, and I fear somewhat the results may not be very helpful for a time. However, it is a start, and will doubtless lead to something definite later.

"In our course we are using Monroe's *Measuring the Results of Teaching* as a text, and for the most part giving to the teachers themselves most of the tests discussed in this book. When the teachers seem familiar with a given test and the use of it, we have them give it in their schools and then we secure their results, discuss them, and ask them to apply whatever suggestions we together may derive from such procedure."

The state normal school at Gunnison has established a School Service Bureau which is serving the county superintendents and teachers of the state by supplying them with the testing material. In doing so, this normal school is following a practice which has been found exceedingly worth while in other states.

National Association of Directors of Educational Research

(E. J. ASHBAUGH, *Secretary and Editor*)

Recently in talking with a gentleman who is greatly interested in the problems of educational research, he ventured the suggestion that most of our bureaus of educational research were misnamed. He stated that he thought they should be named bureaus of accounting, since the work which was being done was almost, if not entirely, of an accounting nature rather than research.

We wonder if the statement is true and if there is not a suggestion contained therein which will be beneficial to the members of our organization. Let each director of an organized bureau of educational research mentally review his work during the past year and consider whether or not it has been research or merely accounting. Has *status* alone been revealed? Have data been gathered, tabulated, and presented without careful analysis of the causal factors and the setting up of a definite remedial program?

We feel that the work of bureaus of educational research is most vital. We believe that the opportunities are practically limitless. We believe, also, that the present is critical and that the future of the movement for organized bureaus rests in the hands of those who are now directing such work. A few cities have established such bureaus and discontinued them because the school administration did not feel that the benefits derived were sufficient to justify their continuance. At this stage of the development of the movement such failure is almost a calamity. Superintendents and teachers may fail and the schools will continue. The educational world is accustomed to such failures. But we must not fail. We must not only discover status but also must reveal tendencies and causes. We must offer to the classroom teacher and the supervisory force definite constructive help in the solution of the problems which are revealed by measurement. We must offer the administrators well-conceived programs based upon definite data which can be presented to school boards and the public. We must keep the scientific attitude but speak in terms of common knowledge. We must attack minute problems and carefully control conditions in our experiments, but keep ever in mind that teachers and the public want results in terms of the larger educational progress.

Unless bureaus of research can modify practice; unless they can offer constructive help all along the line; unless they can make teachers, supervisors, and administrators feel that their work is indispensable, they will fail.

But we shall not fail.

The applications for membership of four persons are before the executive committee at this writing and others are going through the preliminary stage. Interest in the work of the association is certainly growing. Does any member know of a bureau anywhere which is not represented by membership? If so, he will please write to the secretary at once.

LIST OF MEMBERS OF THE NATIONAL ASSOCIATION OF
EDUCATIONAL RESEARCH

HONORARY MEMBERS

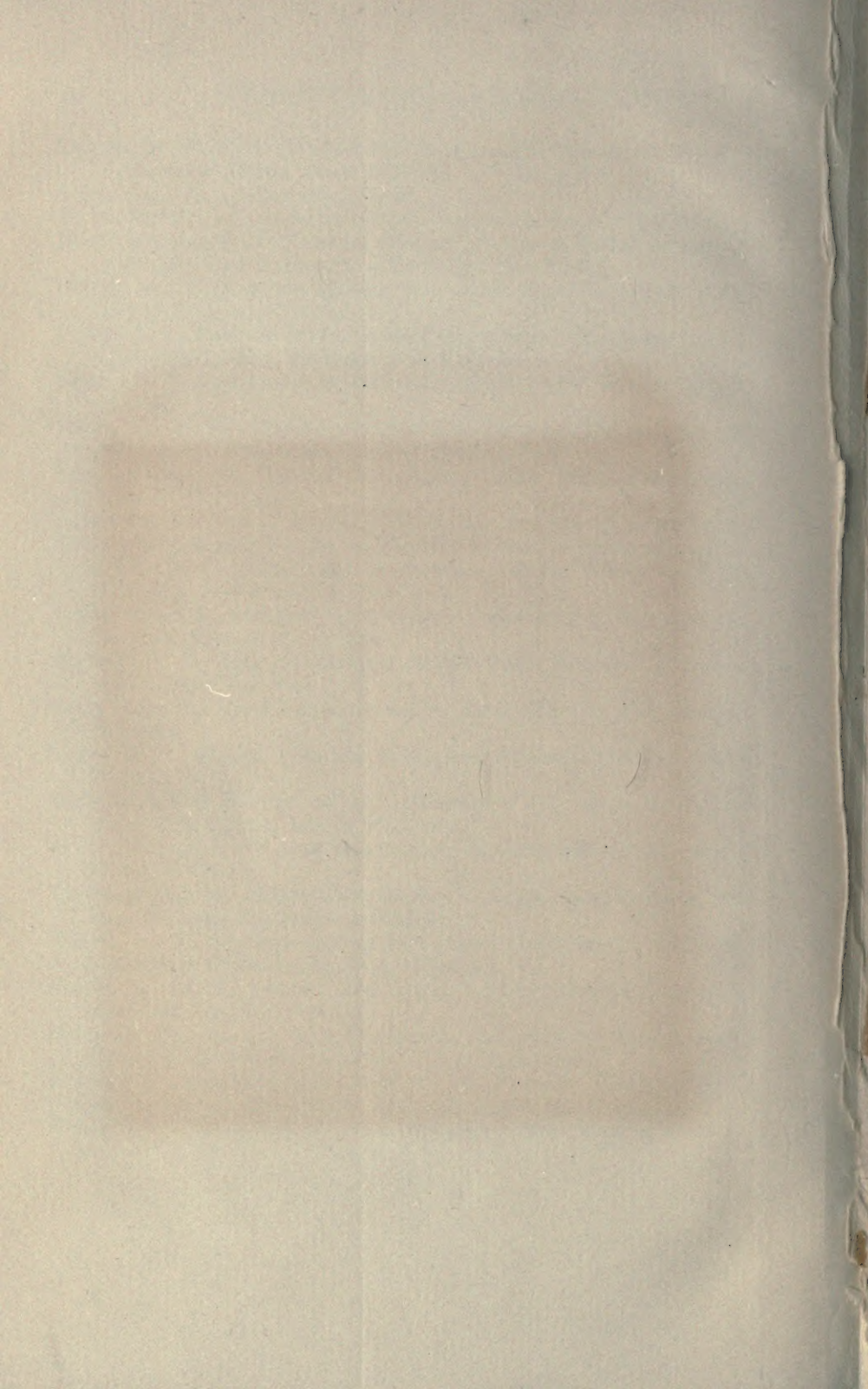
- Chadsey, Dr. C. E., Dean College of Education, University of Illinois, Urbana, Illinois.
Hanus, Dr. Paul H., Harvard College, Cambridge, Massachusetts.
Judd, Dr. Chas. H., University of Chicago, Chicago, Illinois.
Rice, Hon. J. M., Germantown, Pennsylvania.
Russell, Dr. William F., Dean School of Education, State University of Iowa, Iowa City, Iowa.
Thorndike, Dr. E. L., Teachers College, Columbia University, New York City, New York.

REGULAR MEMBERS

- Alexander, Dr. Carter, Assistant State Superintendent, Madison, Wisconsin.
Anderson, H. W., 50 Broadway, Detroit, Michigan.
Ashbaugh, Dr. E. J., Educational Service, Extension Division, State University of Iowa, Iowa City, Iowa.
Averill, William A., Phi Gamma Delta Club, 34 W. 44th Street, New York City, New York.
Ayres, Dr. Leonard P., Director Division of Education, Russell Sage Foundation, New York City, New York.
Ballou, Frank W., Assistant Superintendent of Schools, Mason Street, Boston, Massachusetts.
Barthelmess, Harriet M., Department of Educational Investigation, 34 Normal School Building, Huntington Avenue, Boston, Massachusetts.
Brainerd, Mrs. Margaret, Director of Educational Research, Martin's Ferry, Ohio.
Branson, E. P., Department of Research, Long Beach, California.
Bright, Ira J., Superintendent of Schools, Leavenworth, Kansas.
Brown, Miss Emma M., Director Measurements and Standards, Denver Public Schools, Denver, Colorado.
Brown, H. A., President Oshkosh Normal School, Oshkosh, Wisconsin.
Brueckner, L. J., City Normal School, Detroit, Michigan.
Buckingham, Dr. B. R., Director Bureau of Educational Research, University of Illinois, Urbana, Illinois.
Buckner, Dr. C. A., Bureau of School Service, University of Kansas, Lawrence, Kansas.
Burgess, Mrs. Mary Ayres, Division of Education, Russell Sage Foundation, New York City, New York.
Burgess, W. Randolph, Division of Education, Russell Sage Foundation, New York City, New York.
Carney, C. S., care of Scovell, Wellington & Company, 110 State Street, Boston, Massachusetts.
Courtis, S. A., Director Department of Educational Research, 82 Eliot Street, Detroit, Michigan.

- Crane, Fred D., County Superintendent, Mason City, Iowa.
- Dalman, Murray A., Director of Educational Research, Board of Education, Indianapolis, Indiana.
- Dawson, Charles D., Assistant Superintendent, Grand Rapids, Michigan.
- De Voss, J. C., Assistant Bureau Standards and Measurements, State Normal School, Emporia, Kansas.
- Dickson, Dr. Virgil E., Director of Research and Guidance, Public Schools, Oakland, California.
- Elliott, Dr. Chas. H., Rutgers College, New Brunswick, New Jersey.
- Everly, L. L., Assistant Superintendent and Director of Research, St. Paul, Minnesota.
- Flemming, Mrs. Cecile White, State Supervisor Educational Measurements, State Department of Education, Madison, Wisconsin.
- Fordyce, Dr. Chas., Dean Teachers College, Director Bureau Educational Measurement, University of Nebraska, Lincoln, Nebraska.
- Gray, Dr. W. S., University of Chicago, Chicago, Illinois.
- Greene, Dr. Harry A., Assistant in Educational Service, Extension Division, State University of Iowa, Iowa City, Iowa.
- Guy, J. Freeman, Director Research and Measurement, Pittsburgh, Pennsylvania.
- Haggerty, Dr. M. E., Director Bureau of Co-operative Research, University of Minnesota, Minneapolis, Minnesota.
- Harlan, Charles L., Department of Educational Measurement and Standards, Lewiston State Normal School, Lewiston, Idaho.
- Hebden, Edwin, Director Bureau Statistics and Research, 714 Euclid Avenue, Roland Park P. O., Baltimore, Maryland.
- Henry, Miss Mary Bess, Director of Research, Santa Ana, California.
- Hill, Dr. David Spence, President University of New Mexico, Albuquerque, New Mexico.
- Horn, Dr. Ernest, Director University Elementary School, University of Iowa, Iowa City, Iowa.
- Kallom, Arthur W., Bureau of Investigation and Measurement, 34 Normal School Building, Huntington Avenue, Boston, Massachusetts.
- Kelly, Dr. F. J., Dean School of Education, Director of Bureau of School Service, University of Kansas, Lawrence, Kansas.
- Kelley, Dr. Truman Lee, Teachers College, Columbia University, New York City, New York.
- Kirby, Thomas J., Assistant Director of Research, Kansas City, Missouri.
- Lane, Robert H., Assistant Superintendent of Schools, Los Angeles, California.
- Lefler, M. C., Director of Measurements and Research, Lincoln, Nebraska.
- McCall, Dr. William A., Teachers College, Columbia University, New York City, New York.
- Mallory, Miss Clara, 1016 Liberty Street, Beaumont, Texas.
- Melcher, George, Director Bureau of Research and Efficiency, Library Building, Kansas City, Missouri.

- Monroe, Dr. Walter S., Assistant Director Bureau of Educational Research, University of Illinois, Urbana, Illinois.
- Moore, Supt. M. E., Beaumont, Texas.
- Myers, Miss Anna G., Assistant Director of Research, Kansas City, Missouri.
- Nifenecker, Eugene A., Assistant Director Division of Reference and Research, 500 Park Avenue, New York City, New York.
- Norton, John K., Assistant Director Educational Research, Oakland Public Schools, Oakland, California.
- O'Hern, J. P., Assistant Superintendent, Department of Public Instruction, Municipal Building, Rochester, New York.
- Packer, P. C., Department of Research, Detroit Public Schools, Detroit, Michigan.
- Phelen, Dr. W. W., Director School of Education and Measurement and Efficiency, University of Oklahoma, Norman, Oklahoma.
- Race, Henrietta V., Director Psychological Clinic, Board of Education, Louisville, Kentucky.
- Richardson, Arthur H., Russell Sage Foundation, New York City, New York.
- Richardson, J. Wylie, Educational Research Director, Hibbing, Minnesota.
- Rugg, Dr. H. O., Educational Psychologist, Lincoln School, Columbia University, New York City, New York.
- Sexton, Elmer K., Assistant Superintendent Department of Reference and Research, Newark, New Jersey.
- Shapleigh, F. E., Research Secretary, Buffalo Public Educational Association, Buffalo, New York.
- Smith, Leon O., Director Educational Research, 508 City Hall, Omaha, Nebraska.
- Staker, M. P., Director Bureau of Educational Research, Aberdeen, South Dakota.
- Sutherland, A. H., Director of the Department of Psychology and Educational Research, Los Angeles, California.
- Sutton, Clarence W., Director Reference and Research, Board of Education, Cleveland, Ohio.
- Talbert, Wilford E., Pacific Coast Bureau of Employment Research, 455 Flood Building, San Francisco, California.
- Theisen, Dr. W. W., Supervisor of Educational Measurements, State Department of Education, Madison, Wisconsin.
- Trabue, Dr. M. R., Director Educational Service, Columbia University, New York City, New York.
- Updegraff, Dr. Harlan, Director Bureau Educational Measurements, University of Pennsylvania, Philadelphia, Pennsylvania.
- Welles, J. B., Assistant Director Reference and Research, Cleveland, Ohio.
- Whipple, Dr. G. M., University of Michigan, Ann Arbor, Michigan.
- Woody, Dr. Clifford, University of Washington, Seattle, Washington.



Journal of Educational Research
Vol. No. 1, 1920

NON-CIRCULATING

Journal of Educational Research
Vol. 1, 1920. Jan. - May.

NON-CIRCULATING

